Aus dem Pathologischen Institut der

Ludwig-Maximilians-Universität München

Direktor: Prof. Dr. med. Thomas Kirchner

Clonal dynamics and tumor cell heterogeneity in colorectal cancer

---

Dissertation zum Erwerb des

Doktorgrades der Naturwissenschaften (Dr. rer. nat.)

an der Medizinischen Fakultät

der Ludwig-Maximilians-Universität München

vorgelegt von

Sebastian Lamprecht

aus Linz

2018

Gedruckt mit der Genehmigung der Medizinischen Fakultät der Ludwig-Maximilians-Universität München

Betreuer:	Prof. Dr. rer. nat. Andreas Jung
Zweitgutachter:	Prof. Dr. rer. Nat. Roland Kappler
Dekan:	Prof. Dr. med. dent. Reinhard Hickel

Tag der mündlichen Prüfung: 14.05.2018

Meiner Familie

EIDESSTATTLICHE VERSICHERUNG

Ich erkläre hiermit an Eides statt, dass ich die vorliegende Dissertation mit dem Thema

„Clonal dynamics and tumor cell heterogeneity in colorectal cancer“

selbständig verfasst, mich außer der angegebenen keiner weiteren Hilfsmittel bedient und alle Erkenntnisse, die aus dem Schrifttum ganz oder annähernd übernommen sind, als solche kenntlich gemacht und nach ihrer Herkunft unter Bezeichnung der Fundstelle einzeln nachgewiesen habe.

Ich erkläre des Weiteren, dass die hier vorgelegte Dissertation nicht in gleicher oder in ähnlicher Form bei einer anderen Stelle zur Erlangung eines akademischen Grades eingereicht wurde.

München, den 22.05.2018

Sebastian Lamprecht

PUBLICATIONS

Parts of this thesis have been published in:

- **Lamprecht S**, Schmidt EM, Blaj C, Hermeking H, Jung A, Kirchner T, Horst D. Multicolor lineage tracing reveals clonal architecture and dynamics in colon cancer. Nature Communications 2017; 8: 1406.

- **Lamprecht S**, Kaller M, Schmidt EM, Blaj C, Hermeking H, Grünewald T G, Jung A, Kirchner T, Horst D. PBX3 is part of an EMT regulatory network in colorectal cancer and indicates poor outcome. Clinical Cancer Research 2018; 24: 1974-1986.

In addition, I made contributions to the following publications that are not further described in this work:

- Blaj C, Schmidt EM, **Lamprecht S**, Hermeking H, Jung A, Kirchner T, Horst D. In colorectal cancer, high MAPK activity induces EMT, marks progenitor cells, and remains regulated, irrespective of RAS mutations. Cancer Res. 2017 77(7): 1763-1774

- Blaj C, Bringmann A, Schmidt EM, Urbischek M, **Lamprecht S**, Fröhlich T, Arnold G, Krebs S, Blum H, Hermeking H, Jung A, Kirchner T, Horst D. ADNP is a therapeutically inducible repressor of WNT signaling in colorectal cancer. Clin Cancer Res. 2017 23(11): 2769-2780.

- Woischke C, Blaj C, Schmidt EM, **Lamprecht S**, Engel J, Hermeking H, Kirchner T, Horst D. CYB5R1 links epithelial-mesenchymal transition and poor prognosis in colorectal cancer. Oncotarget. 2016 7(21): 31350-60.

TABLE OF CONTENTS

1. Introduction

1.1. Colorectal Cancer

With around 14.1 million new incidences and 8.2 million deaths worldwide cancer represents a major cause of death throughout the world.
The ongoing growth and aging of the population will increase this liability in both more and less developed countries. Especially lifestyle behaviors like smoking, poor diet, physical inactivity, and reproductive changes – all factors known to increase the risk of cancer - have further raised this burden mainly in less economically developed countries (Botteri et al., 2008; Giovannucci, 2002; Karimi et al., 2014).
Among cancer, colorectal cancer (CRC) serves as a crucial reason of morbidity and mortality worldwide representing over 9% of all cancer cases resulting in 700 thousand deaths in the year 2012. This depicts CRC as the third most common cancer worldwide and the fourth most common cause of cancer related death (Jemal et al., 2009; Torre et al., 2015; World Cancer Research Fund, 2007).
CRC develops in a multistep process, that is characterized by an accumulation of epigenetic and genetic changes leading to the inactivation of tumor suppressive mechanisms and the upregulation of oncogenic pathways, transforming normal glandular epithelium into invasively growing adenocarcinomas. This highly dynamic process is referred to as tumorigenesis. Fearon and Vogelstein first defined the stages of this process in the classic adenoma-carcinoma sequence model where they suggested that only a limited number of genetic alterations drive the formation of CRC (Fearon and Vogelstein, 1990).
By now, studies have revealed mutations in about 67 genes in a colon cancer genome, of those, a subgroup of twelve genes were identified to be predominantly involved in cancer formation. Like other typical solid tumors, CRC requires two to eight of these mutations to develop over time (Sjöblom et al., 2006; Vogelstein et al., 2013). This accumulation of mutations enables

normal tissue to acquire several specific capacities that have been designated as the hallmarks of cancer (Figure 1) (Hanahan and Weinberg, 2000, 2011).

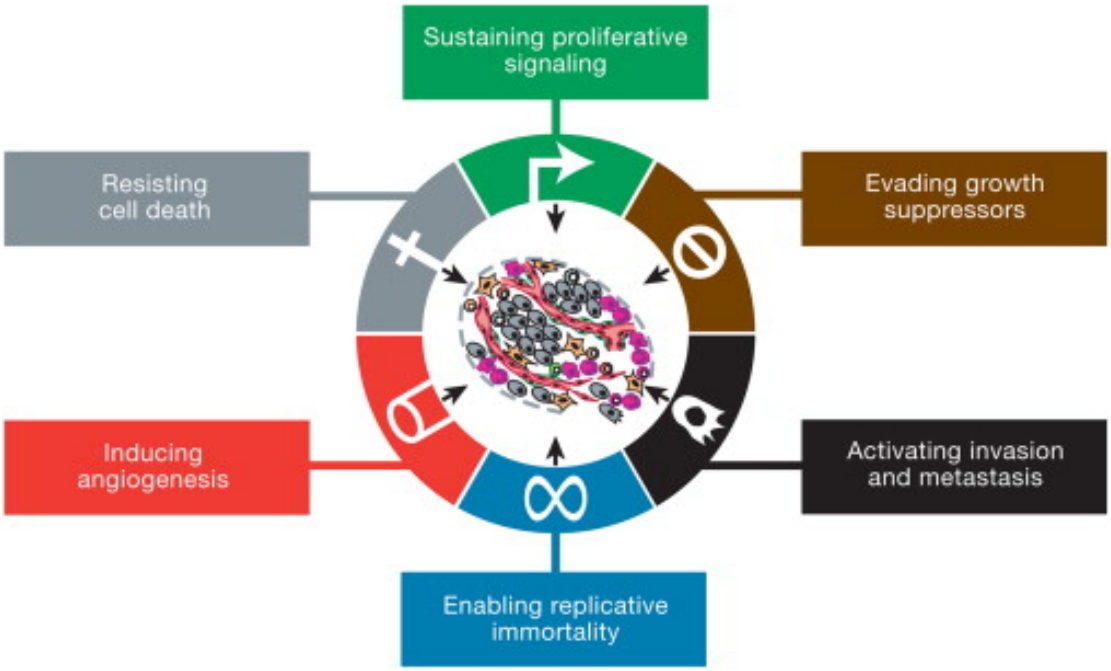

Figure 1: The hallmarks of cancer. The six capacities that enable invasive tumor growth and metastatic propagation. Figure from (Hanahan and Weinberg, 2011).

The relevance of genetic aberrations in tumorigenesis can be observed in small single crypt adenomas; the earliest lesions of the adenoma-carcinoma sequence. These small adenomas typically harbor mutations in APC or β-catenin, resulting in the activation of the WNT pathway (Vazquez et al., 2008; Vogelstein et al., 1988). These mutations provide selective growth advantage to the altered intestinal epithelial cells over surrounding cells – therefor named "driver mutations" whereas a "passenger mutations" does not confer selective growth advantage to cells in which they occur (Vogelstein et al., 2013).

Consecutive alterations in genes like *KRAS* or *TP53* may arise, leading to aberrant activation of MAPK signaling and therefore promote the clonal progression to cancer (Nosho et al., 2008; Samowitz et al., 2005; Vazquez et al., 2008; Vogelstein et al., 2013). Furthermore, this progression can also include mutations in genes like *SMAD2*, *SMAD4*, *RUNX3*, and *TSP1*, leading to a deregulation of crucial signaling pathways including the transforming growth factor - β (TGFB1) pathway (Figure 2) (Lao & Grady, 2011; Grady &

Markowitz 2008; Macías-Silva et al. 1996; Takaku et al. 1998; Wood et al. 2007).)

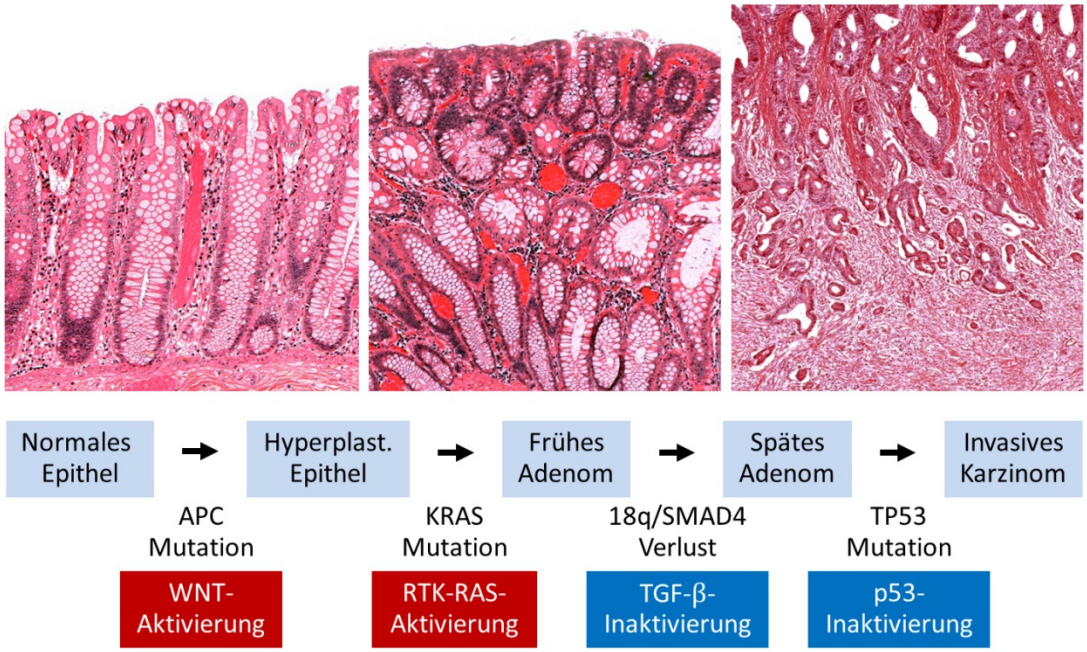

Figure 2: Transformation process from normal colon epithelium to an invasive colorectal carcinoma. Upper panels: H&E staining of normal colon epithelium (left), adenoma (middle) and carcinoma (right). Lower panels: Deregulation of crucial signaling pathways accompanying the adenoma–carcinoma sequence. Figure adapted from (Fearon and Vogelstein, 1990).

In general it is considered that despite a large number of different genomic or epigenomic instabilities that occur in CRC, the selective pressure emerging in the clonal evolution is largely the same, resulting in genes that are more frequently altered in CRC compared to other cancer types (Fearon, 2011).

For instance, alterations affecting the WNT/β-catenin pathway occur in about 60% of all colorectal cancer cases, whereas *KRAS or BRAF* are altered in approximately 55-60%, and 30% carry alterations in the TGFB1 signaling pathway (Grady et al., 1998; Jass et al., 2002; Samowitz et al., 2005). However, CRCs are characterized by substantial genotypic and phenotypic heterogeneity and these traits confer a unique peculiarity to each tumor. Each CRC case thus must be considered as an individual disease with specific characteristics (The Cancer Genome Network Atlas, 2012). Therefore, the discovery of new prognostic and predictive molecular biomarkers is required for a better characterization of CRC and the determination of the most appropriate therapy.

1.2. Tumor heterogeneity and cancer stem cells

The idea of tumor heterogeneity has been established over the last few decades, suggesting that cells that make up a tumor display substantial differences in properties like morphology, cell surface markers and genetic alterations. Further complexity to this heterogeneity is added by different extrinsic stimuli that can affect tumor development and progression including immune cell interactions (Cabrera, Hollingsworth, & Hurt, 2015; Campbell et al., 2010; Shah et al., 2009).

Virchow and Cohnheim were the first to postulate the presence of cancer stem cells (CSCs) within such tumor cell heterogeneity (Huntly and Gilliland, 2005). Later this existence was proven by Bonnet and Dick, isolating cells from acute myeloid leukemia (AML) that were capable to induce tumor growth in immune-compromised recipients (Bonnet and Dick, 1997). Furthermore it is well established that heterogeneity within a tumor (intratumor heterogeneity) evolves over time as CSCs grow and differentiate asymmetrically (Bao et al., 2013; Lathia et al., 2011). By now, the existence of CSCs is quite well accepted but their role in various tumors and how they contribute to tumor formation and tumor cell heterogeneity remain the subject of debate and investigation. Still, studies depicting cellular heterogeneity together with the observation that tumors contain both highly tumorigenic as well as non-tumorigenic cells suggested that cancers have an intrinsic hierarchical organization (Cabrera et al., 2015).

Hence, two distinct models, known as the stochastic or clonal evolution (CE) model and the hierarchy or CSC model, may explain the progression and heterogeneity of tumors (Figure 3) (Gerdes et al. 2014; Michor & Polyak 2010).

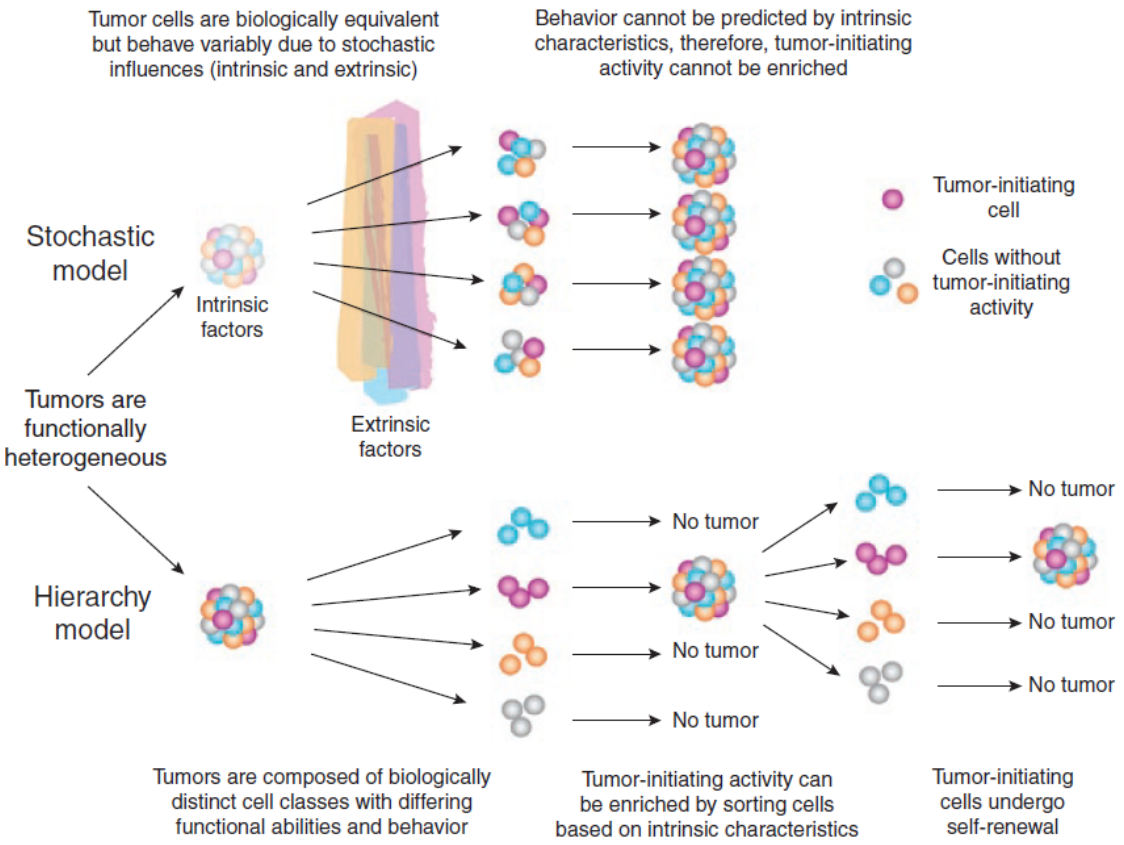

Figure 3: Models of tumor heterogeneity. Schematic illustration of the two models how tumor heterogeneity may arise. Figure from (Dick, 2009)

According to the stochastic model, all malignant cells in principle are biologically equal (Dick, 2009). However, since tumor cells are genetically unstable and their behavior is additionally influenced by intrinsic and extrinsic cues, alterations may accumulate over time and may by chance increase characteristics such as invasiveness, tumor aggressiveness and treatment resistance. Natural selection then drives tumor progression and only few tumor cells may dominate outgrowth within a tumor. Hence according to this model, tumor-initiating cells may not be enriched by sorting cells based on phenotypic characteristics (Gerdes et al., 2014; Michor and Polyak, 2010).

In contrast, the CSC model postulates that tumors are composed of distinct types of cells, each with their own capabilities and phenotypes. Only a specific type of cells, the CSCs, have self-renewing potential and therefore are capable to initiate tumor growth. Moreover, CSCs give rise to non-tumorigenic progeny that may make up most of the tumor mass. Unlike the CE model, the CSC model thus concludes that based on intrinsic characteristics, tumor-

initiating cells can be identified and separated from the non-tumorigenic population (Meacham & Morrison 2013).

As these two models are quite mutually exclusive, an alternative model of reversible cellular plasticity has been proposed, integrating characteristics of both models. In the so called plasticity model, cancer cells are capable to convert between stem cell and differentiated tumor cell states (Plaks et al., 2015). These competing concepts demonstrate that our understanding of cellular heterogeneity in tumors still is incomplete and requires further study.

1.3. Intestinal stem cells and colon cancer stem cells

The healthy human intestine is composed of millions of crypts, containing differentiated cell lineages that reside in distinct functional compartments. Besides stem cells, these include enterocytes, goblet cells, tuft cells and enteroendocrine cells. Stem cells, located in the niche or at the base of the crypt, give rise to more differentiated cells of a transit-amplifying cell lineage. Differentiated colon epithelial cells are then subjected to a massive cellular turnover, being replaced approximately every five days. During this turnover, terminally differentiated cells migrate towards the crypt top and then are shed into the intestinal lumen (Figure 4) (Humphries & Wright 2008; Blanpain et al. 2007; Cernat et al. 2014). Although much effort and interest have been spent, the lineage hierarchy and proliferative potential of intestinal progenitors still remain under debate (Barker et al., 2012). Furthermore, while it is largely acknowledged that the stem cell population occupies the crypt base, the origin of the stem cell niche and the

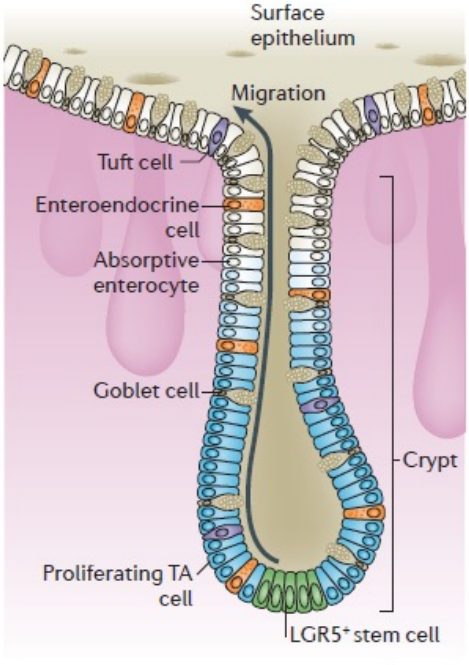

Figure 4: Architecture of the colon epithelium. The colon epithelium is separated in different functional compartments harboring distinct cell lineages. Figure from (Barker, 2014).

rational of their fate are still questionable. To provide further insight, lineage tracing studies have been established. Thereby single cells are marked without changing their features so that the label conveys to the cell's progeny, resulting in clonal expansion of the label. LeBlond and collaborators were the first to utilize lineage tracing, examining the intestinal epithelium by pulse-chase labelling experiments (Bjerknes and Cheng, 1981; Cheng and Leblond, 1974). They hypothesized that all intestinal cell types originate from intestinal stem cells that reside within a small crypt based columnar cell (CBC) population at the crypt bottom. Additionally, radiation-damage studies challenged this theory, proposing that stem cells seem to be located at the forth row from the bottom of the crypt ('+4' cells) (Ponder et al., 1985; Potten et al., 1997). Using newer approaches of transgenic technologies, Barker and colleagues published that the WNT target gene, Leucine repeat-containing G protein coupled receptor 5 (*Lgr5*), is preferentially expressed in cells located at the crypt base (Barker et al., 2007; Van de Wetering et al., 2002). Further studies revealed additional markers, like *Bmi1* (a polycomb RING finger oncogene), *Tert* (telomerase reverse transcriptase), and *Hopx* (HOP homeobox) both of which are preferentially located at position +4 (Montgomery et al. 2011; Takeda et al. 2011; Vermeulen & Snippert 2014). Stem cells thus may comprise different cell populations that express all four of these markers or marker expression may change over time. This led to the question if distinct intestinal stem cell populations may coexist or if these markers may not be defined by a distinct expression pattern (Buczacki et al., 2013; Itzkovitz et al., 2012; Muñoz et al., 2012). To answer this question, an inducible labelling technique, in which a ubiquitously active promoter controls the expression of a Cre transgene, was developed (Lopez-Garcia et al., 2010). Although recombination was restricted to epithelial cells within the niche, only those cells that appeared to attach to the bottom of the crypt survived long term, and subsequently arrange persistent stripes of cells with clonal origin (Figure 5).

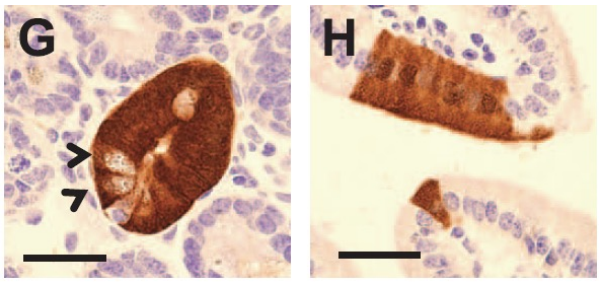

Figure 5: Lineage tracing of the intestine: Longitudinal section of 2-week-old clones. Intestinal cells were genetically labelled with a transgenic mouse model and a ubiquitous promoter. Figure from (Lopez-Garcia et al., 2010).

The power of an inducible genetic-labelling approach then was further improved by Livet and colleagues, who generated a "Brainbow" multicolor reporter transgene that enabled combinatorial expression of different fluorescent proteins (XFPs) in a stochastic manner (Livet et al., 2007). Using three XFPs, red fluorescent protein (RFP), yellow fluorescent protein (YFP) and cyan fluorescent protein (CFP), flanked by three different incompatible *lox* site variants (*loxN*, *lox2272* and *loxP*), independent recombination of those transgene copies allowed generation of distinct color combinations. Cells without recombination of this transgene expressed orange fluorescent protein (OFP) (Figure 6). Studies like these contributed to a better understanding of the architecture of normal colonic mucosa.

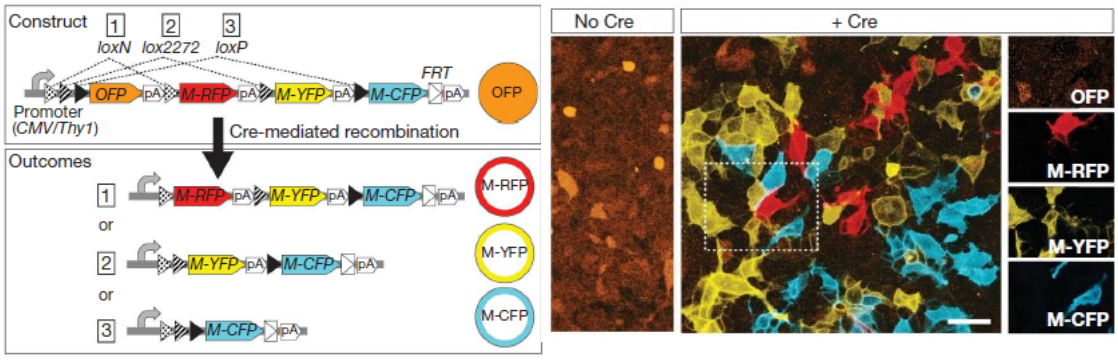

Figure 6: Brainbow 1.1 construct: Left panel: Construct using three different *lox* sites, generating three recombination possibilities. Right panel: Cells carrying this construct. Cre recombination caused expression of M-RFP, M-YFP, or M-CFP. Figure modified from (Livet et al., 2007)

However, in regards to CRC much less is known since these tumors do not preserve the architecture of normal colonic crypts but instead form masses with varying degrees of morphologically disarrayed epithelial glands (Bosman et al., 2010). However, colon cancers do not appear to be completely unorganized. Gradients of less differentiated tumor cells at the leading tumor

edge to glandular differentiated tumor cells in the tumor center can be observed in many cases, and mimic the polarity of normal colonic crypts to varying extent (Brabletz et al., 2001; Cernat et al., 2014). However, compared to normal colonic crypts, such gradients in colon cancer are not situated within stereotypical morphological units and some colon cancers even lack differentiation gradients.

Colon cancer cell subpopulations with distinct phenotypes and degrees of differentiation may have different functions. For example, tumor initiating potential has been attributed to a high WNT and MAPK pathway activity (Vermeulen et al., 2010; Wang et al., 2010). In well-differentiated colon cancers, such tumor cells are frequently located close to the infiltrative tumor edge, leading to the hypothesis that colon cancer stem cells reside at this location (Brabletz et al., 2005). However, defining colon cancer stem cells through tumor-initiating potential, the current “gold standard”, may have certain limitations and cannot always be generalized (Horst et al., 2012; Kreso and Dick, 2014). Moreover, it has been questioned whether the position of a cell within the cellular hierarchy of a growing tumor is adequately reflected by tumor-initiating potential (Clevers, 2011). Therefore, from these data, the role of distinct tumor cell phenotypes for the dynamics of clonal expansion in colon cancer has remained unclear.

Using similar lineage tracing tools as the ones applied for studies in normal colonic mucosa, clonal dynamics in genetically engineered mouse tumor models have been analyzed (Driessens et al., 2012; Schepers et al., 2012). Moreover, current studies demonstrated clonal outgrowth from colon cancer cells with high MAPK activity or expression of the WNT target gene *LGR5*, and thus provided direct evidence for a cellular hierarchy emanating from these tumor cell subsets *in vivo* (Blaj et al., 2017; Shimokawa et al., 2017). Despite these data, the role of distinct tumor cell phenotypes in colon cancer still remains a subject of debate and needs to be further evaluated.

1.4. Epithelial-mesenchymal transition

Further complexity in regard to tumor cell heterogeneity is added by a fundamental mechanism described as epithelial-mesenchymal transition (EMT) Hereby, epithelial cells lose their distinctive features and become more migratory (Figure 7). EMT is thought to be central for cancer invasion and considered hallmark of cancer progression (Hanahan and Weinberg, 2011). This process was initially observed by Elizabeth Hay as transformations in cell structure, that facilitates cells to move into the inner of an embryo and contribute to internal organ development (Hay, 1995; Nieto, 2013; Thiery et al., 2009). The capability of epithelial cells to acquire a mesenchymal state and to transform back again in a reverse procedure, called mesenchymal-epithelial transition (MET), suggested plasticity of epithelial cell phenotypes (Tam and Weinberg, 2013).

In the process of EMT, epithelial cells lose their polarity, remodulate their cytoskeleton which leads to an increased motility that enables the transformation into an invasive phenotype (Thiery and Sleeman, 2006; Thiery et al., 2009). This is mainly induced by micro environmental signals, leading to the activation of transcription factors (TFs) that collaborate with epigenetic regulators and therefor change the translation of proteins affecting cell polarity, cytoskeleton structure, and extracellular matrix degradation including the repression of key epithelial differentiation genes (Tam and Weinberg, 2013).

The crucial event during EMT is considered to be downregulation of the cell adhesion molecule E-cadherin (encoded by *CDH1*). Additional suppression of genes encoding claudins and occludins lead to destabilization of apical tight junctions and loss in epithelial barrier function (Huang et al., 2012b; Peinado et al., 2007). Furthermore, downregulation of E-cadherin promotes the expression of genes that increase mesenchymal adhesion such as neural cadherin (N-cadherin) and other markers like LAMC2 or Vimentin (Nieto et al., 2016; Wheelock et al., 2008).

On the molecular level, the decrease of E-cadherin expression is mediated by transcription factors impairing the *CDH1* promoter. They can be classified into two groups that either exhibit direct or indirect effects on E-cadherin

expression. SNAIL ZEB1, the Krüppel-like factor KLF8 and E47 directly bind and repress the activity of the *CDH1* promoter, whereas factors such as TWIST, the forkhead-box protein FoxC2 and E2.2 repress *CDH1* transcription indirectly by interacting with miRNAs (Figure 7) (Batlle et al., 2000; Cano et al., 2000; Gonzalez and Medici, 2014; Huang et al., 2012a; Mani et al., 2007; Oliver E. Owen, 2002; Peinado et al., 2007; Pérez-Moreno et al., 2001; Sánchez-Tilló et al., 2012; Thiery et al., 2009; Wang et al., 2007; Yang et al., 2004).

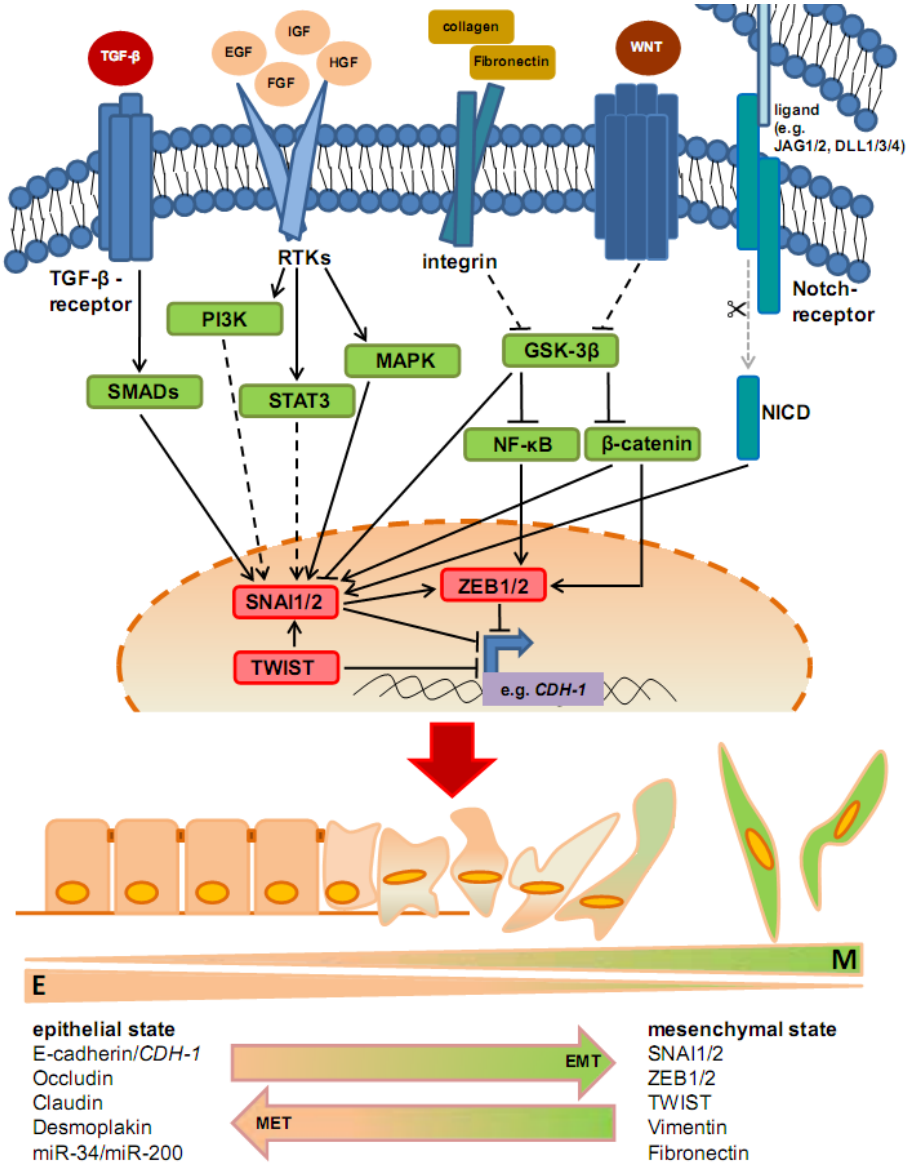

Figure 7: Molecular mechanisms in the EMT network. In the process of EMT epithelial cells (orange, E) transform to mesenchymal cells (green, M). The reverse process termed as MET (mesenchymal-epithelial transition). Figure from (Hahn and Hermeking, 2014).

Furthermore, it has been shown that mechanisms comparable to EMT can also occur as response to injury or during tumorigenesis, in the formation of metastasis, and in lesions implicating organ degeneration, such as fibrosis (Puisieux, Brabletz, & Caramel, 2014; Arnoux, Nassour, L'Helgoualch, Hipskind, & Savagner, 2008; Iwano et al., 2002; Peinado et al., 2007).
In colorectal cancers, EMT features have been described to occur mainly at the invasive edge where tumor cells directly interact with surrounding stromal tissue, causing intratumoral gradients of EMT to MET from tumor edge towards the differentiated, epithelial tumor center. In addition, the mesenchymal phenotype at the invasive edge of the tumor is often accompanied by an accumulation of nuclear β-catenin, indicating an active WNT pathway (Brabletz, 2012; Brabletz et al., 2005). Moreover, it has been shown that the WNT pathway directly induces SNAIL1 activity by phosphorylation by GSK3β (Yook et al., 2006; Zhou et al., 2004), linking WNT signaling and EMT. Cooperation of other signaling pathways, like Notch, MAPK, PI3K-AKT or TGFβ also force the initiation and progression of EMT, highlighting the complexity of the EMT signaling network and the importance in understanding its fundamental mechanisms to eventually develop potential therapeutic strategies.

1.5. WNT signaling pathway

WNT signaling is a fundamental pathway found in metazoan animals. WNT originates from the Drosophila polarity gene wingless and the vertebrate analog, *integrated* (Wodarz and Nusse, 1998). Signaling through the WNT pathway is one of the fundamental mechanisms affecting cell polarity, proliferation and embryonic development (Logan and Nusse, 2004). Hence, alterations in this pathway often cause human birth defects, neurological disorders and various cancers (Clevers, 2006; Clevers and Nusse, 2012; De Ferrari and Moon, 2006).
WNTs are secreted proteins that are encoded by 19 different genes in the human genome. These proteins are characterized by a similar sequence

pattern that encodes for proteins with a distinct cysteine pattern and further conserved residues, instead of functional features. With more than 15 different receptors and co-receptors the WNT pathway accounts for a signaling cascade that is extremely complex and often regulated by feed-back control (Miller, 2002).

The crucial event of the canonical or WNT/β-catenin signaling pathway is the translocation of the adherens junction associated protein β-catenin into the nucleus (Figure 8). If WNT is absent, cytoplasmic β-catenin is continuously phosphorylated from the β-catenin destruction complex, that includes the scaffolding protein Axin, the tumor suppressor *adenomatous polyposis coli* gene product (APC), casein kinase 1 (CK1), and glycogen synthase kinase 3 β (GSK-3β) (Amit et al., 2002; Liu et al., 2002; Yanagawa et al., 2002; Yost et al., 1996). Together, this inhibits that β-catenin translocates into the nucleus and represses the expression of WNT target genes by the T cell factor/lymphoid enhancer factor (TCF/LEF) (Aberle et al., 1997).

If a WNT ligand is present, the low-density lipoprotein receptor related proteins 5/6 (LRP5/6), form complexes with WNT-bound Frizzled, leading to the activation of the scaffolding protein Dishevelled (Dvl). A following displacement of GSK-3β from APC/Axin results in the recruitment of the destruction complex to receptors, thus inhibiting Axin-mediated β-catenin phosphorylation. This increase of β-catenin levels leads to its nuclear accumulation, replaces Groucho from TCF/LEF and activates the transcription of WNT target genes (Clevers, 2006; Gammons et al., 2016; Li et al., 2012).

The importance of altered WNT/β-catenin signaling in cancer, particularly in CRC is well documented. Although WNT is constitutively active, due to inactivating mutations in APC or activating β-catenin mutations, WNT signaling in CRC still remains regulated on high levels, leading to different tumor cell populations with low or high WNT activity (Horst et al., 2012). It has been shown that colon cancer cell subpopulations harboring high levels of WNT activity are typically associated with mesenchymal characteristics such as putative cancer stem cell traits and a marker expression profile that is linked to tumor invasion (Brabletz et al., 2005). These cells reside at the infiltrative tumor edge where they can invade the surrounding tissue. On the

contrary, cells with low WNT signaling commonly occupy the center of the tumor (Cernat et al., 2014; Kirchner and Brabletz, 2000). Due to these findings, high WNT signaling activity is assumed to be a driving force of colon cancer invasion and progression, making it an attractive potential target for therapeutic intervention (Kahn, 2014).

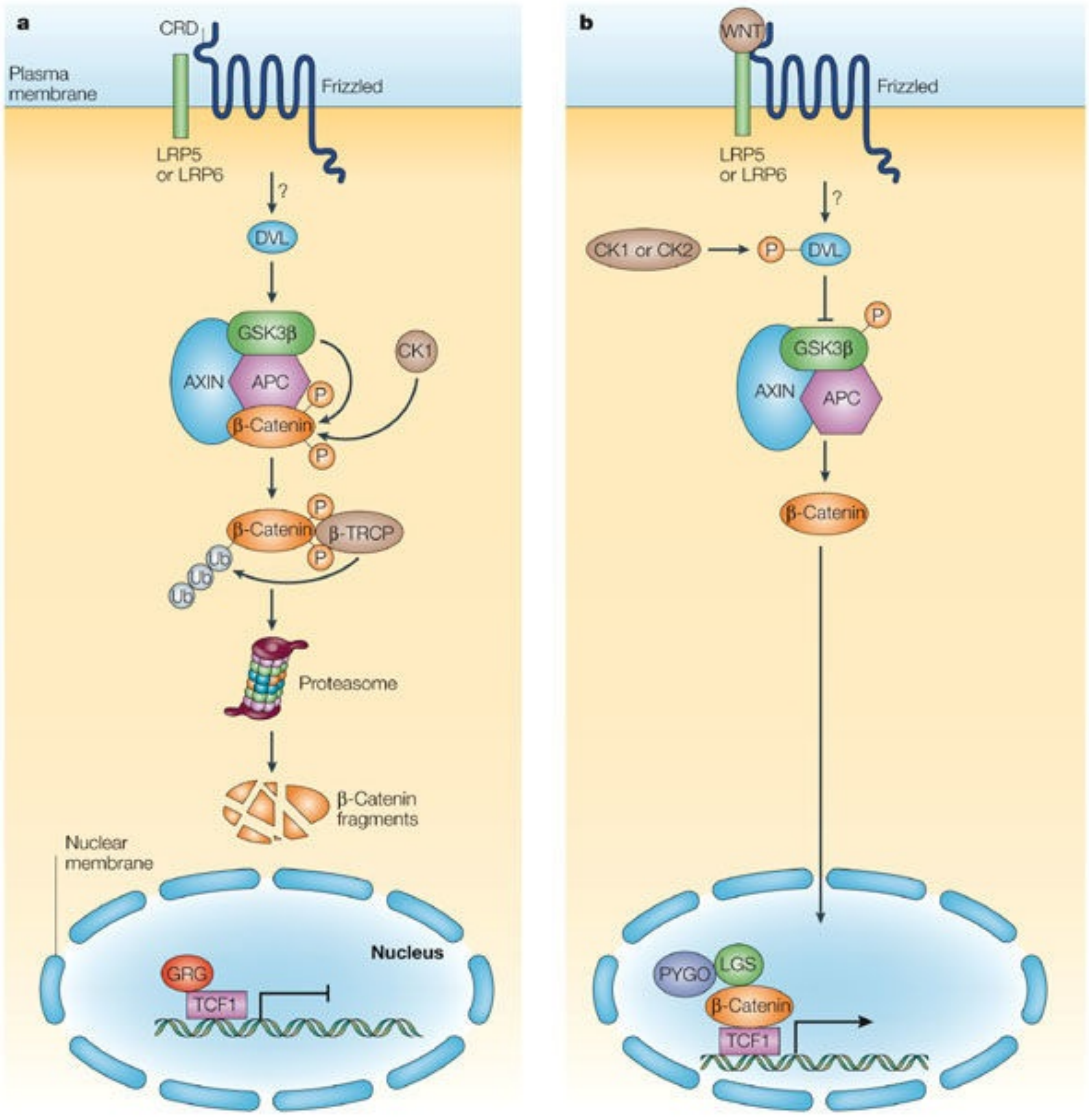

Figure 8: Summary of the WNT signaling pathway. (A) In the absence of WNT, β-catenin binds to the deconstruction complex that leads its degradation. **(B)** WNTs bind to their receptors resulting in an inactivation of GSK3β and further to an activation of WNT target genes by β-catenin. Figure from (Staal and Clevers, 2005).

1.6. Pre-B-cell leukemia transcription factor 3

Distinct degrees of EMT and WNT signaling in colon cancer exemplify that colorectal cancers are composed of phenotypically different cell subpopulations in the same genetic background. In the second part of this thesis, we searched for unknown factors that are associated with tumor cell heterogeneity. In this context, we identified high expression of the TALE transcription factor Pre-B-cell leukemia homeobox transcription factor 3 (PBX3) in tumor cells undergoing EMT. PBX3 belongs to a transcription factor family that is described to facilitate tumor growth. Enhanced expression of PBX3 is associated with tumor growth and progression in various cancer types like ovarian cancer, melanoma and prostate cancer (H.-B. Han et al., 2014; Crijns et al., 2007; Kikugawa et al., 2006; Shiraishi et al., 2007). Although some research has been done on the other PBX homologues, the biological function of PBX3 remains unknown (Y. Li et al., 2014; Monica, Galili, Nourse, Saltman, & Cleary, 1991). Recently, it has been shown in prostate and colorectal cancer that PBX3 was upregulated and the expression was mediated by androgen through micro RNA let-7d (Han et al., 2012; Ramberg et al., 2011). Furthermore, other studies revealed a correlation of high PBX3 expression with the invasiveness of CRC cells and an association with metastasis (Han et al., 2014). However, the contribution of PBX3 to human colorectal cancer and its functional role in tumor progression has remained unclear.

2. Aims of the thesis

This thesis had the following aims:

I) Determination of the clonal architecture and dynamics in colon cancer

II) Identification and characterization of PBX3 expression in colorectal cancer

3. Materials

3.1. Chemicals and reagents

Compound	Supplier
4-Hydroxytamoxifen ≥70% Z isomer (remainder primarily E-isomer)	Sigma-Aldrich, St. Louis, MO, USA
4x Laemmli Sample Buffer	Bio-Rad, München, Germany
All purpose Hi-Lo DNA Marker	Bionexus, Netanya, Israel
Agarose Biozym LE	Biozym Scientific GmbH, Hessisch Oldendorf, Germany
Ampicillin sodium salt	Sigma-Aldrich, St. Louis, MO, USA
APS (ammonium peroxodisulfate)	Carl Roth GmbH, Karlsruhe, Germany
beta-Mercaptoethanol	Carl Roth GmbH, Karlsruhe, Germany
Biofreeze	Biochrom GmbH, Berlin, Germany
Blasticidin	Carl Roth GmbH, Karlsruhe, Germany
BSA (Albumin Faktor V)	Carl Roth GmbH, Karlsruhe, Germany
BSA Standard Set	Bio-Rad, Munich, Germany
Bovine serum albumin 25% (BSA)	Thermo Fisher Scientific Inc., Waltham, MA, USA
Chloramphenicol	Carl Roth GmbH, Karlsruhe, Germany
Chlorophorm	Sigma-Aldrich, St. Louis, MO, USA
cOmplete mini protease inhibitor cocktail	Roche Diagnostics GmbH, Mannheim, Germany
DAPI (2-(4-amidinophenyl)-6-indolecarbamidine dihydrochloride)	Carl Roth GmbH, Karlsruhe, Germany
dNTPs (deoxynucleotides triphosphate)	Thermo Fisher Scientific Inc., Waltham, MA, USA
DMEM	Biochrom GmbH, Berlin, Germany
DMEM/F-12 with GlutaMAX medium	Thermo Fisher Scientific Inc., Waltham, MA, USA
DMSO (dimethyl-sulfoxide)	Carl Roth GmbH, Karlsruhe, Germany
dNTP Mix	Thermo Fisher Scientific Inc., Waltham, MA, USA
Doxycycline hyclate	Sigma-Aldrich, St. Louis, MO, USA
ECL/HRP substrate	Immobilon, Merck Millipore, Billerica, MA, USA
EGF Recombinant Human Protein	Thermo Fisher Scientific Inc., Waltham, MA, USA
Ethidiumbromidlösung 1%	Carl Roth GmbH, Karlsruhe, Germany
Fast-Media Amp Agar	InvivoGen, San Diego, CA, USA
Fast SYBR Green Master Mix	Applied Biosystems, Foster City, CA, USA
FBS (fetal bovine serum)	Biochrom GmbH, Berlin, Germany
FGF-Basic Recombinant Human Protein	Thermo Fisher Scientific Inc., Waltham, MA, USA

Compound	Supplier
FuGENE6 Transfection Reagent	Promega, Madison, WI, USA
GE Healthcare Chromatography Paper	Schubert & Weiss Omnilab GmbH & Co. KG, München, Germany
Hi-Di Formamide	Applied Biosystems, Foster City, CA, USA
HiPerFect Transfection Reagent	Qiagen GmbH, Hilden, Germany
ImmEdge Hydrophobic Barrier PAP Pen	Biozol GmbH, Eching, Germany
Immobilon-P Transfer Membrane	Immobilon, Merck Millipore, Billerica, MA, USA
LB medium (Luria/Miller)	Carl Roth GmbH, Karlsruhe, Germany
LipoD293	Tebu-Bio, Le Perray En Yvelines, France
Matrigel	Corning, New York City, NY, USA
NP40 Substitute	Sigma-Aldrich, St. Louis, MO, USA
Opti-MEM	Thermo Fisher Scientific Inc., Waltham, MA, USA
PageRuler Plus Prestained Protein Ladder	Fermentas GmbH, St. Leon-Rot, Germany
Paraformaldehyde	Carl Roth GmbH, Karlsruhe, Germany
Penicillin/Streptomycin	Biochrom GmbH, Berlin, Germany
PhosSTOP Phosphatase Inhibitor Ccocktail	Roche Diagnostics GmbH, Mannheim, Germany
ProLong Gold Antifade	Invitrogen GmbH, Karlsruhe, Germany
Propidium iodide	Sigma-Aldrich, St. Louis, MO, USA
Puromycin dihydrochloride	Merck KGaA, Darmstadt, Germany
Rotiphorese Gel 30 (37,5:1)	Carl Roth GmbH, Karlsruhe, Germany
SDS (sodium dodecyl sulfate)	Carl Roth GmbH, Karlsruhe, Germany
Skim milk powder	Sigma-Aldrich, St. Louis, MO, USA
sunflower oil	Sigma-Aldrich, St. Louis, MO, USA
StemPro hESC Supplement	Thermo Fisher Scientific Inc., Waltham, MA, USA
Tamoxifen free base	Sigma-Aldrich, St.Louis, MO, USA
Target Retrieval Solution 6 (TRS)	Agilent Technologies, Inc., Santa Clara, CA, USA
Temed (tetramethylethylendiamin,1,2-bis (dimethylamino) -ethan)	Carl Roth GmbH, Karlsruhe, Germany
Triton X 100	Carl Roth GmbH, Karlsruhe, Germany
Trizol Reagent	Invitrogen GmbH, Karlsruhe, Germany
Trypsin/ EDTA solution	Biochrom GmbH, Berlin, Germany
Tween 20	Sigma-Aldrich, St. Louis, MO, USA
WNT3a	R&D Systems, Minneapolis, MN, USA

3.2. Enzymes

Enzyme	**Supplier**
DNAse I (RNAse-free)	Sigma-Aldrich, St. Louis, MO, USA
FastAP Thermosensitive Alkaline Phosphatase	Thermo Fisher Scientific Inc., Waltham, MA, USA
FIREPol DNA Polymerase	Solis BioDyne, Tartu, Estonia
Klenow Fragment	Thermo Fisher Scientific Inc., Waltham, MA, USA
Pfu Polymerase (recombinant)	Thermo Fisher Scientific Inc., Waltham, MA, USA
Proteinase K	Thermo Fisher Scientific Inc., Waltham, MA, USA
T4 DNA polynucleotide kinase	Thermo Fisher Scientific Inc., Waltham, MA, USA

3.3. Kits

Kit	**Supplier**
DyeEx 2.0 Spin Kit	QIAGEN GmbH, Hilden, Germany
Pure Yield Plasmid Midiprep System	Promega GmbH, Mannheim, Germany
QIAprep Spin Miniprep Kit	QIAGEN GmbH, Hilden, Germany
Quantitect Reverse Transcription Kit	QIAGEN GmbH, Hilden, Germany
QuikChange II XL Site-Directed Mutagenesis Kit	Stratagene, Agilent Technologies GmbH & Co.KG, Waldbronn, Germany
RevertAid H Minus First Strand cDNA Synthesis Kit	Thermo Fisher Scientific, Inc., Waltham, MA, USA
Vectastain ABC Kit Universal	Vectorlabs, Burlingame, CA, USA
Wizard SV Gel and PCR Clean-Up System	Promega GmbH, Mannheim, Germany

3.4. Antibodies

3.4.1. Primary antibodies

Antigen	Source/Clone	Application	Supplier
α-tubulin	Mouse/ DM 1A	WB	Sigma-Aldrich, St. Louis, MO, USA
Active- β-catenin	Mouse/ 8E7	WB	Merck KGaA, Darmstadt, Germany
β-catenin	Mouse/ clone 14	IHC	Ventana Medical Systems, Oro Valley, AZ, USA
β-catenin	Mouse	WB	BD PharMingen, Heidelberg, Germany
BrdU	Mouse	IHC	Santa Cruz Biotechnology, Inc., Heidelberg, Germany
CK20	Goat	IHC, IF	Santa Cruz Biotechnology, Inc., Heidelberg, Germany
Cre	Rabbit	WB	Cell Signaling Technology, Inc., Danvers, MA, USA
FLAG	Rabbit	IHC, IF	Cell Signaling Technology, Inc., Danvers, MA, USA
FRA-1	Rabbit	IHC, IF	Abcam, Cambridge, UK
E-cadherin	Mouse/ 24E10	WB, IF	Cell Signaling Technology, Inc., Danvers, MA, USA
GFP	Mouse/ 4B10	IF	Cell Signaling Technology, Inc., Danvers, MA, USA
GLUT1	Mouse/ A-4	IHC, IF	Santa Cruz Biotechnology, Inc., Heidelberg, Germany
Ki67	Rabbit	IHC	Cell Signaling Technology, Inc., Danvers, MA, USA
Laminin5γ2	Mouse/ D4B5	IHC	Merck Millipore KGaA, Darmstadt, Germany
PBX3	Mouse/ M01	WB, IHC, IF	Abnova, Taipei, Taiwan
Snail	Rabbit	WB, IF	Cell Signaling Technology, Inc., Danvers, MA, USA
V5	Goat	IHC, IF	Abcam, Cambridge, UK
VSV	Rabbit/ P5D4	IHC, IF	Abcam, Cambridge, UK
Vimentin	Rabbit	WB	Cell Signaling Technology, Inc., Danvers, MA, USA
ZEB1	Rabbit/ H-102	WB	Santa Cruz Biotechnology, Inc., Heidelberg, Germany

WB: Western blot analysis, IF: immunofluorescence, IHC: immunohistochemistry

3.4.2. Secondary antibodies

Name	Source/Clone	Application	Supplier
Anti-Rabbit HRP	Goat	WB	Sigma-Aldrich, St. Louis, MO, USA
Anti-Mouse HRP	Donkey	WB	Promega GmbH, Mannheim, Germany
Anti-Goat Alexa Fluor 488	Donkey	IF	Thermo Fisher Scientific, Inc., Waltham, MA, USA
Anti-Mouse Alexa Fluor 568	Donkey	IF	Abcam, Cambridge, UK
Anti-Rabbit Alexa Fluor 568	Donkey	IF	Abcam, Cambridge, UK
Phalloidin-Alexa-568	-	IF	Thermo Fisher Scientific, Inc., Waltham, MA, USA

WB: Western blot analysis, IF: immunofluorescence

3.5. DNA constructs and oligonucleotides

3.5.1. Template vectors

Name	Insert	Reference
CMV-Brainbow	Kusabira Orange, mCherry, mEYFP, M-mCerulean	(Livet et al., 2007)
pBV-luc		(He et al., 1999)
pBV-PBX3	PBX3 - promoter	
PBX3 geneblock	human *PBX3* 3'UTR from -2674 to -1185 relative to the translational start site	
pCMV-dR8.91	Gag-Pol	(Zufferey et al., 1997)
pLenti CMV rtTA3G Blast	Reverse Tetracycline transactivator 3G	Dominic Esposito
pLenti CMVTRE3G eGFP Puro	Enhance Green Fluorescent Protein	Eric Campeau
pMD2.G	VSV G	Didier Trono
pRTR		(Jackstadt et al., 2013)
pRTR-Snail	human Snail	(Siemens et al., 2011)
pRTR-ZEB1	human ZEB1	
pGL3-control-MCS		(Kaller et al., 2011; Welch et al., 2007)
pGL3-PBX3 wt	human *PBX3* 3'UTR	
pGL3- PBX3 mut	human *PBX3* 3'UTR	
pRL	Renilla	(Pillai, 2005)

3.5.2. Primers

Gene	Sequence (5’ – 3’)	Purpose
Axin 2 fwd	AGGCCAGTGAGTTGGTTGTC	qPCR
Axin 2 rev	CATCCTCCCAGATCTCCTCA	qPCR
CTNNB1 fwd	AGCTGACCAGCTCTCTCTTCA	qPCR
CTNNB1 rev	CCAATATCAAGTCCAAGATCAGC	qPCR
CDH1 fwd	ATCCAAAGCCTCAGGTCATA	qPCR
CDH1 rev	CAGCAAGAGCAGCAGAAT	qPCR
CreERT2 fwd	ATCCACCTGATGGCCAAG	qPCR
CreERT2 rev	GCTCCATGCCTTTGTTACTCA	qPCR
GAPDH fwd	ACCACAGTCCATGCCATCAC	qPCR
GAPDH rev	TCCACCACCCTGTTGCTGTA	qPCR
LGR5 fwd	TACCCACAGAAGCTCTGCAGAATT	qPCR
LGR5 rev	TGTTCAGGGCCAAGGTCATG	qPCR
NKD-1 fwd	TCACTCCAAGCCGGCCGCC	qPCR
NKD-1 rev	TCCCGGGTGCTTCGGCCTATG	qPCR
PBX3 fwd	GCCTTGGAGGAAATTCACTG	qPCR
PBX3 rev	AGATGGAGTTGTTGCGTCCT	qPCR
PBX3 3'UTR fwd	TAAGAATTCGATCAGAGACTGGTAGCATCG	PCR
PBX3 3'UTR rev	ATAACCGGTAATCATGAAAGCAAAAAGTTTATTC	PCR
PBX3 3'UTR mut fwd	GAAATATACAGTACTGAAAAGTCAAATCTGAATGCATCACAA TTAGTCGCTGCTTTT	PCR
PBX3 3'UTR mut rev	AAAAGCAGCGACTAATTGTGATGCATTCAGATTTGACTTTTC AGTACTGTATATTTC	PCR
PBX3 5'UTR fwd	CTCTAAGCGCTTTGCGATTG	PCR
PBX3 5'UTR rev	AGCATCCTGGATTGATCGTC	PCR
pri-miR-200c fwd	CTTAAAGCCCCTTCGTCTCC	qPCR
pri-miR-200c rev	AGGGGTGAAGGTCAGAGGTT	qPCR
Snail fwd	GCACATCCGAAGCCACAC	qPCR
Snail rev	GGAGAAGGTCCGAGCACA	qPCR
Vimentin fwd	TACAGGAAGCTGCTGGAAGG	qPCR
Vimentin rev	ACCAGAGGGAGTGAATCCAG	qPCR
ZEB1 fwd	TCAAAAGGAAGTCAATGGACAA	qPCR
ZEB1 rev	GTGCAGGAGGGACCTCTTTA	qPCR

fwd = forward, rev = reverse, PCR = polymerase chain reaction, qPCR = quantitative (real time) reverse transcription PCR,

3.5.3. siRNAs and MicroRNA mimics

Scramble siRNA, β-Cat siRNA1 and β-Cat siRNA2 were purchased from QIAGEN. pre-miR-200c, siRNAs against PBX3 and ZEB1 were obtained from Ambion.

3.6. Buffers and solutions

50x TAE buffer

- 40 mM Tris Base
- 20 mM acetic acid
- 1 mM EDTA pH 8.0
- ad 1 liter ddH_2O

10x Vogelstein‘ PCR buffer:

- 166 mM NH_4SO_4
- 670 mM Tris (pH 8.8)
- 67 mM $MgCl_2$
- 100 mM β-mercaptoethanol

RIPA buffer (for protein lysates):

- 1% NP40
- 0.5% sodium deoxycholate
- 0.1% SDS
- 150 mM NaCl
- 50 mM TrisHCl (pH 8.0)
- ad 100 ml ddH_2O

10x Tris-glycine-SDS running buffer:

- 1.92 M glycine
- 250 mM Tris base
- 1% SDS
- pH 8.3-8.7
- ad 5 liters ddH_2O

10x Transfer buffer:

- 1.92 M glycine
- 250 mM Tris base
- 1% SDS
- pH 8.3-8.7
- 25% Methanol
- ad 1 liter ddH_2O

10x TBS:

- 20 mM Tris base
- 150 mM NaCl
- ad 5 liters ddH_2O

1x TBST (10l):

- 20 mM Tris base
- 150 mM NaCl
- 0.1% Tween 20
- ad 10 liters ddH_2O

Cell culture medium completed:

- 500 ml DMEM
- 10% FCS
- 1% Penicillin/Streptomycin

Freezing medium:

- FCS
- 10% DMSO

TRIS buffer:

- 1 M Tris base
- pH 7.4
- ad 1 liter ddH_2O

3.7. Laboratory equipment

Device	**Supplier**
5415R table-top centrifuge	Eppendorf AG, Hamburg, Germany
ABI 3130 genetic analyzer capillary sequencer	Applied Biosystems, Foster City, USA
Axioplan 2	Carl Zeiss GmbH, Oberkochen, Germany
Axiovert 25 microscope	Carl Zeiss GmbH, Oberkochen, Germany
BD Accuri C6 Flow Cytometer Instrument	Accuri, Erembodegem, Belgium
BD FACSAria III	Becton Dickinson, Franklin Lakes, NJ, USA
CF40 Imager	Kodak, Rochester, New York, USA
Falcons, dishes and cell culture materials	Schubert & Weiss OMNILAB GmbH & Co. KG
Fisherbrand FT-20E/365 transilluminator	Fisher Scientific GmbH, Schwerte, Germany
Forma scientific CO_2 water jacketed incubator	Thermo Fisher Scientific, Inc., Waltham, MA, USA
GeneAmp PCR System 9700	Applied Biosystems, Foster City, USA
HERACell 240i Co2 Incubator	Thermo Fisher Scientific, Inc., Waltham, MA, USA
Herasafe 2020 safety cabinet	Thermo Fisher Scientific, Inc., Waltham, MA, USA
HTU SONI130	G. Heinemann Ultraschall- und Labortechnik, Schwäbisch Gmünd, Germany
Light Cycler 480 II	Hoffmann-La Roche AG, Basel, Schweiz
LSM 700	Carl Zeiss GmbH, Oberkochen, Germany
ME2CNT membrane pump	Vacuubrand GmbH & CO KG, Wertheim, Germany
Megafuge 1.0R	Heraeus; Thermo Fisher Scientific, Inc., Waltham, MA, USA
Mini-PROTEAN-electrophoresis system	Bio-Rad, München, Germany
MultiImage Light Cabinet	Alpha Innotech, Johannesburg, South Africa
ND 1000 NanoDrop Spectrophotometer	NanoDrop products, Wilmington, DE, USA
Neubauer counting chamber	Carl Roth GmbH & Co, Karlsruhe, Germany
Orion II luminometer	Berthold Technologies GmbH & Co. KG, Bad Wildbad, Germany
Peqpower	Peqlab Biotechnologie GmbH, Erlangen, Germany
PerfectBlue SEDEC 'Semi-Dry' blotting system	Peqlab Biotechnologie GmbH, Erlangen, Germany
Primo Vert microscope	Carl Zeiss GmbH, Oberkochen, Germany
T100 Thermo Cycler	Bio-Rad, München, Germany
Thermoblock comfort	Eppendorf AG, Hamburg, Germany
Varioskan Flash Multimode Reader	Thermo Scientific, Inc., Waltham, MA, USA
Waterbath	Memmert GmbH, Schwabach, Germany

4. Methods

4.1. Bacterial cell culture

Standard cloning procedures were carried out using Stbl3 and Dh5α bacterial *E.coli* strains (Invitrogen). Bacteria were cultured at 37°C overnight in LB medium or on LB agar plates to isolate single cell colonies. Resistant clones bearing a resistance cassette were selected by addition of ampicillin (100 µg/ml).

Plasmid transformation into bacteria was achieved by adding 100 ng of plasmid DNA into competent *E.coli* followed by a 30 minutes incubation step on ice. After 90 seconds of heat shock at 42°C cells were placed on ice for another two minutes and then plated on ampicillin containing LB-agar plates at 37°C overnight. To extract plasmid DNA, single clones of the bacterial cultures were cultivated in ampicillin containing LB-medium for 8-12 hours and QIAprep Spin Miniprep Kit (Qiagen) was used.

To identify and determine the orientation of bacterial clones harboring a vector insert, single clones were collected and a colony PCR was performed. Finally, PCR products were visualized on a 1% agarose gel. For amplification of plasmid DNA, 50 ng DNA was transferred into a 40 µl reaction mix containing 1 µl dNTPs, 4 µl 10x PCR buffer, 2 µl DMSO, 1 µl Pfu DNA polymerase and 1 µl of forward and reverse primers each. PCR cycling conditions were as exemplarily given: three minutes at 95°C, followed by 30 cycles of 95°C for 30 seconds, 60°C for 90 seconds, 72°C for 30 seconds and a termination cycle for 5 minutes at 72°C.

4.2. Mammalian cell culture

4.2.1. Cultivation of human cell lines

Colo320, DLD-1, HEK293, HCT116, LS174T, LoVo, SW620 and SW480 cell lines were obtained from ATCC, SW1222 were a gift from the Ludwig Institute for Cancer Research (New York, USA). LS174T dnTCF4 and DLD-1 dnTCF4 were a kind gift from M. van de Wetering (Hubrecht Institute, Utrecht; (Van de Wetering et al., 2002)). Cells were cultured in Dulbecco's Modified Eagle's Medium (DMEM) completed with 10% fetal bovine serum (FBS), 100 U/ml penicillin, and 0.1 mg/ml streptomycin at 37°C and 5% CO2. For induction experiments, doxycycline (DOX) was always used at a concentration of 100 ng/ml.

4.2.2. Generation of DNA constructs

The pRTR-ZEB1-VSV vector was generated by an excision of the ZEB1 cDNA from pcDNA-His-MaxC-ZEB1 (a kind gift from Janet E. Mertz, McArdle Laboratory for Cancer Research, University of Wisconsin School of Medicine and Public Health; (Ellis-Connell et al., 2010)) with NotI and XbaI and cloned into pUC19-SfiI. The DNA sequence encoding the N-terminal His-tag was excised with NotI and BamHI and replaced by complementary oligonucleotides encoding a VSV-tag. The VSV-ZEB1 encoding sequence was then excised with SfiI and cloned into pRTR (Jackstadt et al., 2013).
The pRTR-SNAIL-VSV vector has been described recently (Siemens et al., 2011). Hereby cells were transfected with the pRTR expression vector using Fugene6. Afterwards positive cells were selected with 2 µg/ml Puromycin for two weeks. Transfection efficiency was checked by adding DOX at a final concentration of 100 ng/ml and counting GFP-positive cells.
Cloning of the PBX3 3'-UTRs with putative miR-200c binding sites was done by PCR amplification of the human PBX3 mRNA from SW480 cells using Pfu Polymerase. The obtained PBX3 mRNA sequences were inserted into pGL3-

control vector. Mutation of the miR-200 binding sequences of human *PBX3* 3'UTRs was done with the QuickChange Mutagenesis Kit (Stratagene) and checked by sequencing.

For analysis of WNT/β-catenin activity within the *PBX3* promotor, 2500 bp 5' of the *PBX3* transcription start site were obtained by PCR amplification from a human BAC clone (Life technologies) using Pfu Polymerase. The obtained promoter sequences were inserted into pBV luc-control vector. Synthetic DNA sequences (IDT) were used to replace TCF4 binding sites by mutated sites and checked by sequencing.

For the inducible pLenti TetO-CreERT2 expression vector, we PCR amplified CreERT2 from pCAG-CreERT2 (Diaz Jr et al., 2012), and inserted it between BamH1 and Xba1 restriction sites of pLenti CMVTRE3G eGFP Puro (a gift from Eric Campeau), replacing eGFP by CreERT2. For the Cre sensitive recombination vector pLenti Multicolor, we first PCR-amplified expression cassettes for Kusabira orange, mCherry, and EYFP from CMV-Brainbow 1.1 M (Gatenby et al., 2009), and EBFP2 from pEBFP2-Nuc (Matsuda and Cepko, 2007), using primers that omitted membrane or nuclear localization signals, respectively. Amplicons then were inserted into EcoRV sites of pcDNA3.1(+) (Invitrogen), and the 3' ends of mCherry, EYFP, and EBFP2 were replaced from BsrG1 to Not1 restriction sites by synthetic sequences that added FLAG, V5, or VSV tags, respectively. Kusabira orange and tagged fluorescent color coding genes then were sequentially inserted into a plasmid with synthetic paired loxN, lox2272, and loxP sites. The whole expression cassette then was inserted between Age1 and Sal1 sites of pLenti PGK-GFP (a gift from Didier Trono), replacing GFP. Finally, the PGK promoter was replaced by an EF1α promoter, yielding pLenti Multicolor. To check DNA sequences of all vectors, samples were sent to GATC Biotech for Sanger sequencing.

4.2.3. Lentivirus production

Lentivirus was produced in HEK293 cells that had been seeded 24 hours before by co-transfection of 10 µg of the lentiviral vector with 10 µg pCMV-dR8.91, and 2 µg pMD2.G using 60 µl LipoD293. After 15 minutes of incubation the transfection mix was added drop-wise to unsupplemented DMEM. After 12 hours, the medium was changed with fresh completed DMEM. 36 hours post transfection, the virus containing culture supernatant was collected, and given on pre cultured CRC cell lines for infection. HCT116 and SW1222 cells were sequentially stably transduced with pLenti CMV rtTA3G Blast and pLenti TetO-CreERT2. Cells were then selected with puromycin and blasticidin, and subsequently infected with pLenti Multicolor. After 3-5 days, cells with orange fluorescence were single cell subcloned by limiting dilution and expanded. To test recombination in vitro, CreERT2 expression was induced in cells by 1 µg/ml DOX for 4 days, and recombination was induced with 10 µM 4-hydroxytamoxifen. After 3-10 days, cells were inspected by fluorescence microscopy.

4.2.4. Transfection of siRNAs and plasmids

Transient transfections with Individual synthesized siRNAs were carried out using a transfection mix containing 250 µl Opti-MEM, 6 µl HiPerFect and 6 µM of oligonucleotide obtaining a final concentration of 100 nM, was set up. For pre-miR-200c 30 nM and WNT3a 20ng/ml were used, respectively. After 15 minutes of incubation, the transfection mix was given slowly to the cells. Transfection of reporter plasmids in a 100 mm dish was carried out using 4 µg DNA and 5 µl FuGENE in a 300 µl transfection master mix.

4.3. Tumor xenografts and *in vivo* treatments

Mouse experiments were reviewed and approved by the Regierung von Oberbayern. 10^6 single clone expanded SW1222 or HCT116 colon cancer cells, carrying the multicolor lineage tracing constructs, were suspended in 100 µl of a 1:1 mixture of PBS and growth factor-depleted Matrigel (Corning), and injected subcutaneously into 6-8 week old NOD/SCID mice (NOD.CB17-Prkdcscid, The Jackson Laboratory) for xenograft formation. When tumor diameters reached 7 mm, recombination of pLenti Multicolor transgenes was induced by 1 mg doxycycline p.o. for 3 consecutive days, followed by 3 mg tamoxifen i.p. (Sigma Aldrich). Mice were sacrificed and tumors were removed at 3, 10, 17, 24, or 31 days after induction. For BrdU tracing, mice were injected once with a 1.25 mg BrdU pulse. At distinct time points mice killed, tumors were removed, fixed in formalin and embedded in paraffin for further analyses.

4.4. RNA Isolation , reverse transcription and qPCR

RNA was collected with Trizol Reagent. Next, cDNA for each probe was synthesized from 500 ng RNA using the Reverse Transcription Kit. qPCR was carried out by using a LightCycler 480 and SYBR Green Master Mix applying 40 cycles of amplification at 95°C (1 sec), 60°C (20 sec), and 72 °C (1 sec). Obtained results were then assigned by normalization to the mRNA levels of the house-keeping gene GAPDH. Primer specificity was checked by recording a melting curve for the PCR products. For sequences of oligonucleotides used as qPCR primers see 3.5.2.

4.5. Protein isolation and Western blot analysis

For Western Blotting, cells were cultured under indicated conditions, harvested and resuspended in RIPA buffer with protease- and phosphatase-inhibitors. The cell lysate was further sonicated for ten seconds and then separated by centrifugation for 15 minutes at 14.000 g at 4°C.
Protein concentrations were measured in a Varioskan Plate Reader using the BSA Standard Set. 50 µg of protein diluted in Laemmli buffer were then denatured for 5 minutes at 95°C, and loaded on a 10% SDS- acrylamide Gel. Gel electrophoresis was carried out at 120 V in an electrophoresis system using a Tris-glycine-SDS running buffer. Transfer of the proteins form the Gel to an Immobilon-P PVDF membrane was done, using a 'Semi-Dry' blotting system at a constant current of 120 mA and 10 V for around 90 minutes. To avoid nonspecific binding of the primary antibodies the membrane was incubated in 5% skim milk/TBS-T for at least 60 minutes. Antibodies were diluted in BSA and applied at 4°C over-night. After extensive washing with TBS-T membranes were incubated in diluted horseradish-peroxidase (HRP)-conjugated antibodies for 60 minutes. For detection of protein bands the membrane was incubated with an ECL/HRP substrate and signals were detected by an Imager (Kodak). Applied Antibodies are listed in the table in chapter 3.4.

4.6. Luciferase assay

HEK293T, Colo320 and SW480 cells were transfected in 24-well plates using 50 ng of a firefly luciferase reporter plasmid, 10 ng of the Renilla reporter plasmid and 20 ng/mL of WNT3a, or 30 nM of pre-miR-200c. 48 hours later, the intensity of the luciferase activity was detected with an Orion II luminometer and further analyzed with the SIMPLICITY software.

4.7. Immunofluorescence and Immunohistochemistry

Immunohistochemical staining was performed using 5µm sections as previously described (Brabletz et al., 2001). Deparaffinization was achieved by xylol and ethanol and antigen retrieval was done in a pH6 buffer for 20 min. Next, the sections were stained with primary antibodies on a Ventana Benchmark XT autostainer with Universal DAB and alkaline phosphatase detection kits.

Stained slides then were inspected by light microscopy for the distribution of each marker-antigen and categorized as negative, polarized if expression gradients from leading tumor edge to tumor center were observed, or diffuse if such gradients were absent. Ki67 proliferation was separately assessed at the leading tumor edge and in the tumor center. Relative BrdU staining intensity was quantified continuously from leading tumor edge to tumor center using ImageJ (NIH).

For immunofluorescent staining, cells were cultivated under respective conditions. Slides or cover-slips were then consecutively stained for 60 minutes with primary antibodies and subsequently secondary Alexa Fluor 405, 488 or 555 conjugated antibodies were used for visualization, and nuclei were counterstained with DAPI. An Alexa Fluor 647-labeled Phalloidin antibody was applied to detect stress fiber forming F-actin. Images were taken on a LSM 700 laser scanning microscope using a Plan Apochromat 20x/0.8 M27 objective and the ZEN 2009 software (Zeiss). Antibodies used are listed un 3.4.

4.8. Clinical case collections

Samples of colorectal cancer patients that underwent intentionally surgical resection between 1994 and 2006 at the LMU were obtained from the archives of the institute of pathology. Follow-up data were documented prospectively by the tumor registry Munich. This collection was assembled respectively to the guidelines of the ethics committee of the Medical Faculty of

the LMU. For the survival collection, inclusion criteria were localized UICC stage II colorectal cancers. Finally, the collection consisted of 244 colorectal cancer samples of which in 52 (21.3%) patients had died of their tumor within the follow-up period. Survival data were censored when case follow-up was discontinued or when patients had died of reasons other than colorectal cancer. For the metastasis collection a case control design was chosen and tumor specimens of 90 patients with right sided colon cancers were included. Half of the patients had colon cancers with synchronous liver metastasis, where metastasis were diagnosed by clinical imaging or liver biopsy. Controls consisted of colon cancer patients without distant metastasis at the time of diagnosis and with a disease-free survival of at least 5 years after primary surgical resection. Cases and controls were matched by tumor grade (according to WHO 2010) and T-category, resulting in 45 matched pairs. Of both collections tissue microarrays (TMAs) were generated with 6 representative 1 mm cores of each case.

4.9. Gene expression datasets and GSEA

For comparative analyses of tumor cell subpopulations, three gene expression data sets derived from colon cancer cells with low and high WNT activity were screened for consistently deregulated genes (Horst et al., 2012; Vermeulen et al., 2010). For analysis of colon cancer samples, microarray data produced on Affymetrix HG-U133Plus2.0 arrays were obtained from the Gene Expression Omnibus (GEO) and normalized with Robust Multi-array Average (RMA) using custom brainarray CDF (v19, ENTREZG) in R, which yielded one optimized probeset per gene, as previously described (Grünewald et al., 2016; Orth et al., 2015; Sahay et al., 2015). Samples used were GSE14333, GSE17536, and GSE39582. Pearson correlations of PBX3 expression and expression of all other genes represented within these datasets were calculated, and genes were ranked accordingly. GSEA analyses then were done using this ranked gene list against curated sets of

EMT core signatures (Anastassiou et al., 2011; Taube et al., 2010). Heatmaps for selected genes were drawn with GENE-E (Broad Institute).

4.10. Panel sequencing

For next-generation panel sequencing, we used the Ion AmpliSeq Cancer Hotspot Panel v2, covering the mutational status of 50 oncogenes and tumor suppressor genes, according to the manufacturers protocol (Life Technologies). 31 days after recombination, individual clones from immunohistochemically stained slides of different SW1222 and HCT116 xenograft tumors were microdissected and 1-5 ng DNA were used as template for library construction. Multiplexed libraries then were sequenced on an Ion Personal Genome Machine (Thermo Fisher). Reads were mapped to human reference genome hg19 and filtered for non-synonymous variants.

4.11. Analysis of clone characteristics

To determine clone sizes, we counted neighboring tumor cells with identical fluorescent colors on confocal images. For each clone, we then determined the positions of each cell C (x_C, y_C), as well as the closest positions of leading tumor edge E (x_E, y_E) and tumor necrosis N (x_N, y_N) using ImageJ (NIH). Using geometric shifting and rotation, we then transformed coordinates so that E' (0, 0) and N' (x_N', y_N') with x_N' = y_N'. The resulting cell positions C' (x_C', y_C') for each clone then were analyzed in Microsoft Excel for linear correlation by t test, the slope of the linear regression m was determined, and the angle θ of the line of best fit with the x-axis was calculated by $\theta = \tan^{-1}(m)$. The angle α of each clone relative to a tangent to the leading tumor edge then resulted from $\alpha = \theta + 45°$.

4.12. Simulation Model

The two-dimensional spatial simulation model was implemented in VBA-Excel (Suppl. Excel Application). In a worksheet “Clones”, simulating 60x60 cells, random numbers from 1 to 3,600 were distributed. These are illustrated in a 60x60 matrix in worksheet “Graphics” with ten different colors, determined by clone number modulo 10. For each simulation cycle for cells at the bottom row, representing the leading tumor edge, each cell content is either copied to neighboring cells on the left or right, simulating lateral expansion for clonal competition, or the cell above, simulating clonal outgrowth towards the tumor center, while this behavior is determined at random. For all remaining cells, contents are copied to the cell above, while this is restricted to every N^{th} row, with N simulating the proliferation gradient from leading tumor edge (bottom row) to tumor center (other rows). Contents of cells that are to be replaced are shifted to cells immediately above, causing loss of “clones” only at the top row of the model, which simulated tumor cells next to tumor necrosis. Frequencies are recorded in worksheet “Numbers” and represented in a diagram in worksheet “Graphics”.

5. Results

5.1. Multicolor lineage tracing reveals structural organization and clonal dynamics in colon cancer

The results presented in this section are part of the publication: **Lamprecht S**, Schmidt EM, Blaj C, Hermeking H, Jung A, Kirchner T, Horst D. Multicolor lineage tracing reveals clonal architecture and dynamics in colon cancer. Nature Communications 2017; 8: 1406.

5.1.1. Differentiation gradients in colon cancer

First we assessed primary colon cancers for the expression of nuclear β-catenin and FRA1 as surrogate markers for high WNT and MAPK signaling (Tetsu and McCormick, 1999; Vial et al., 2003) that were previously linked to tumor initiating potential and colon cancer stem cells. In addition we determined expression of CK20 and GLUT1 that in contrast indicated epithelial cell differentiation and hypoxia, respectively (Bristow and Hill, 2008; Vial et al., 2003). Many colon cancers showed increased nuclear β-catenin and FRA1 expression in tumor cells located at the infiltrative tumor edge, whereas CK20 and GLUT1 were most strongly expressed in the tumor center, often close to necrotic areas, suggesting strong differentiation gradients directed from the tumor edge towards the tumor center. However, a substantial number of colon cancers did not show definite intratumoral differentiation gradients, since they either expressed these markers more randomly throughout the tumor, or were negative for individual markers (Figures 9A and B).

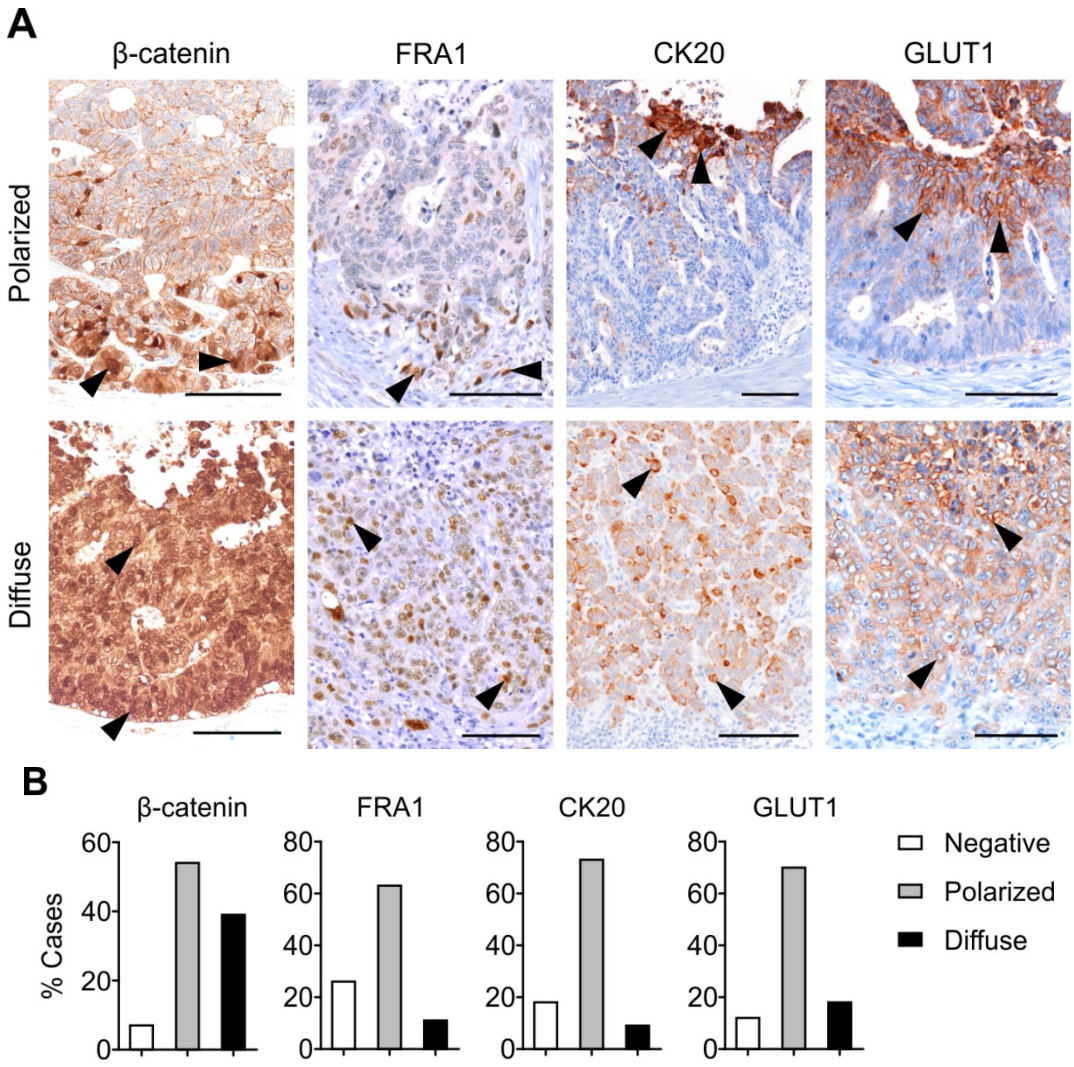

Figure 9: Varying degrees of differentiation gradients in colon cancer. (A) Immunohistochemistry for indicated proteins in representative primary colon cancers used to classify organized and disorganized tumor growth based on presence or absence of strong differentiation gradients. All micrographs show tumors from leading tumor edge (image bottom) to tumor center or central tumor necrosis (image top). Arrowheads indicate positively stained tumor cells. Scale bars 100µm. **(B)** Frequencies of observed marker distributions in colon cancer (n=92).

These findings suggested that colon cancers may be categorized into tumors with polarized or more diffuse expression of differentiation antigens and markers related to colon cancer stem cells (Figure 10).

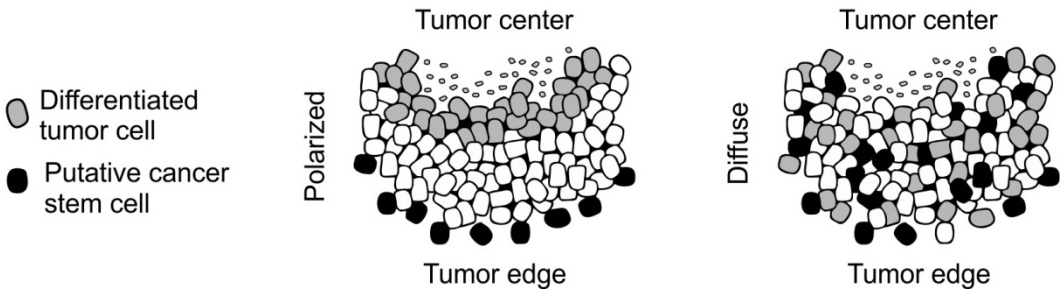

Figure 10: Differentiation gradients in colon cancers. Schematic model for organized (left panel) and disorganized (right panel) colon cancers with and without differentiation gradients, respectively.

Next, we characterized a collection of colon cancer xenografts and found that SW1222 derived tumors showed the same distribution of nuclear β-catenin, FRA1, CK20 and GLUT1 as primary colon cancer cases, while HCT116 colon cancer xenografts showed more diffuse marker expression and lack of differentiation gradients (Figure 11). We therefore used xenografts of these two cell lines as model tumors for the typical spectrum of presence or absence of differentiation gradients that is observed in primary colon cancers.

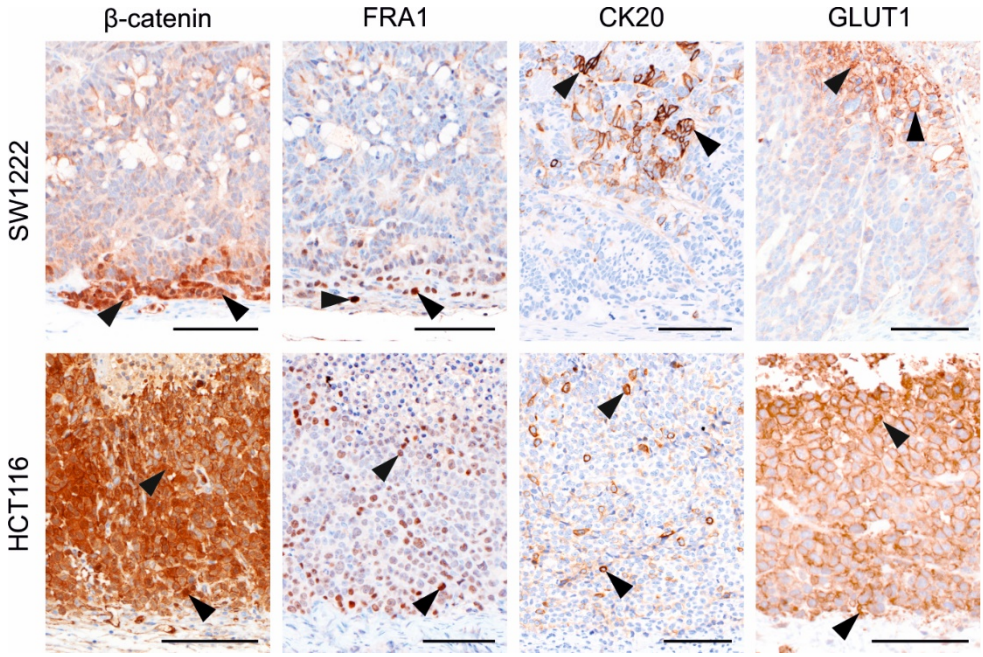

Figure 11: Xenografts of colon cancer cell lines reflect primary colon cancer architecture. Immunohistochemistry for indicated proteins in SW1222 and HCT116 derived xenografts demonstrate presence or absence of strong differentiation gradients, respectively. All micrographs show tumors from leading tumor edge (image bottom) to tumor center or central tumor necrosis (image top). Arrowheads indicate positively stained tumor cells. Scale bars 100 µm.

5.1.2. Multicolor lineage tracing of colon cancer cells *in vivo*

In order to visualize lineage outgrowth in colon cancer, we developed a lentiviral Cre recombinase sensitive reporter system that allows stochastic expression of different fluorescent colors in individual tumor cells. Our system consists of three lentiviral vectors, two of which mediate doxycycline-inducible expression of an estrogen receptor-Cre fusion protein (pLenti rtTA3G and pLenti TetO-CreERT2). Upon Cre recombination the third vector randomly switches from expression of orange to either tagged red, yellow, or blue fluorescence proteins (pLenti Multicolor, Figure 12).

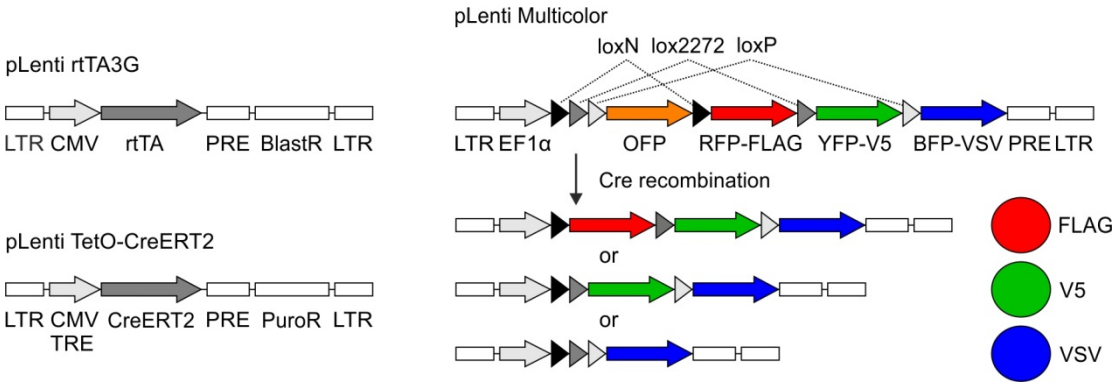

Figure 12: Lentiviral vectors for expression of rtTA (pLenti rtTA3G), doxycycline dependent CreERT2 (pLenti TetO-CreERT2), and the Cre-responsive multicolor transgene (pLenti Multicolor). Upon Cre-recombination, transgene elements flanked by loxN, lox2272, or loxP sites will be removed at random, causing an irreversible switch from expression of orange (OFP) to tagged red (RFP-FLAG), yellow (YFP-V5), or blue fluorescence (BFP-VSV), respectively. LTR, long terminal repeat; TRE, tetracycline response element; BlastR/PuroR, blasticidin and puromycin resistance genes. PRE posttranscriptional regulatory element.

This doxycycline and tamoxifen controlled design was completely devoid of unwanted background recombination (data not shown). We then transduced all three vectors into SW1222 and HCT116 colon cancer cells, expanded single cell clones, and xenografted them into immune compromised NOD/SCID mice (Figure 13A). After xenograft growth, we induced recombination by doxycycline and tamoxifen treatment and analyzed clonal outgrowth over time (Figure 13B).

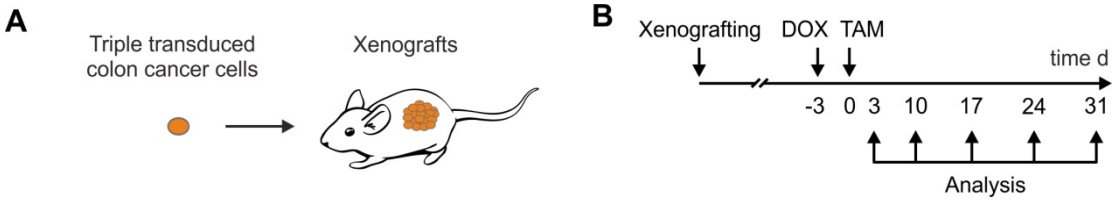

Figure 13: Workflow for Xenograft experiments. (A) Triple transduced colon cancer cells were xenografted into NOD/SCID mice. **(B)** Schedule for doxycycline (DOX) and tamoxifen (TAM) treatment, and tumor harvest after xenografting

Three days after induction of recombination, individual or small clusters of colon cancer cells were randomly labelled by different fluorescent colors in a mosaic pattern throughout SW1222 and HCT116 xenograft tumors. Of note, multiple vector integrations resulted in various combinatorial fluorescent colors (Figure 14).

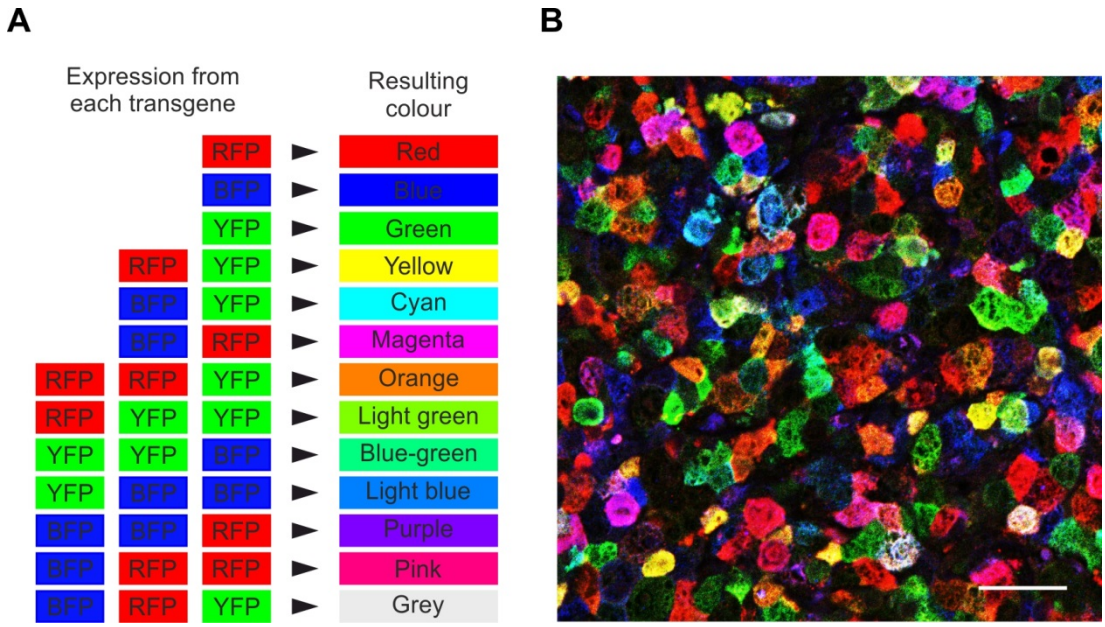

Figure 14: Mixed fluorescent color expression due to multiple vector insertions. (A) Upon insertions of e.g. three copies of pLenti Multicolor, by chance one, two or three inserted transgenes may recombine, resulting in up to 13 different single or mixed fluorescent colors. **(B)** Fluorescent color spectrum in an HCT116 xenograft tumor three days after recombination. Scale bar, 25 µm.

Interestingly, at this early time point after recombination, we already observed loss of few color labelled clones into the central tumor necrosis (Figure 15). Ten days after recombination, single color clones had increased in size, while at 31 days after recombination, large stripe- and wedge-like shaped clones completely extended from the tumor edge to the necrotic tumor center (Figures 15 and 16). Inducible multicolor labeling thus allowed us to monitor clonal outgrowth within human colon

cancer *in vivo*, and suggested clonal expansion along axes from the tumor edge towards the tumor center.

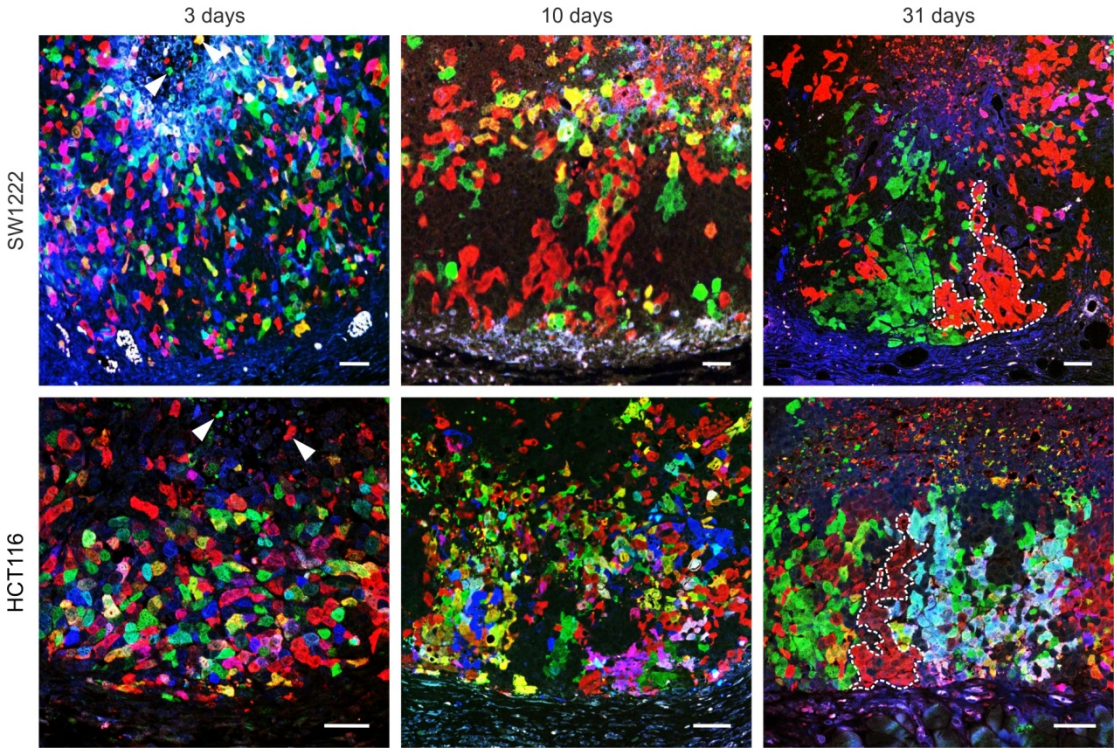

Figure 15: Stochastic clone formation *in vivo*. Confocal immune fluorescence for RFP-FLAG (red), YFP-V5 (green) and BFP-VSV (blue) in indicated xenografts (n ≥ 3) at indicated time points after tamoxifen induced multicolor labeling. Fluorescent images show xenograft tumors from leading tumor edge (image bottom) to central tumor necrosis (image top). Scale bars, 50 µm

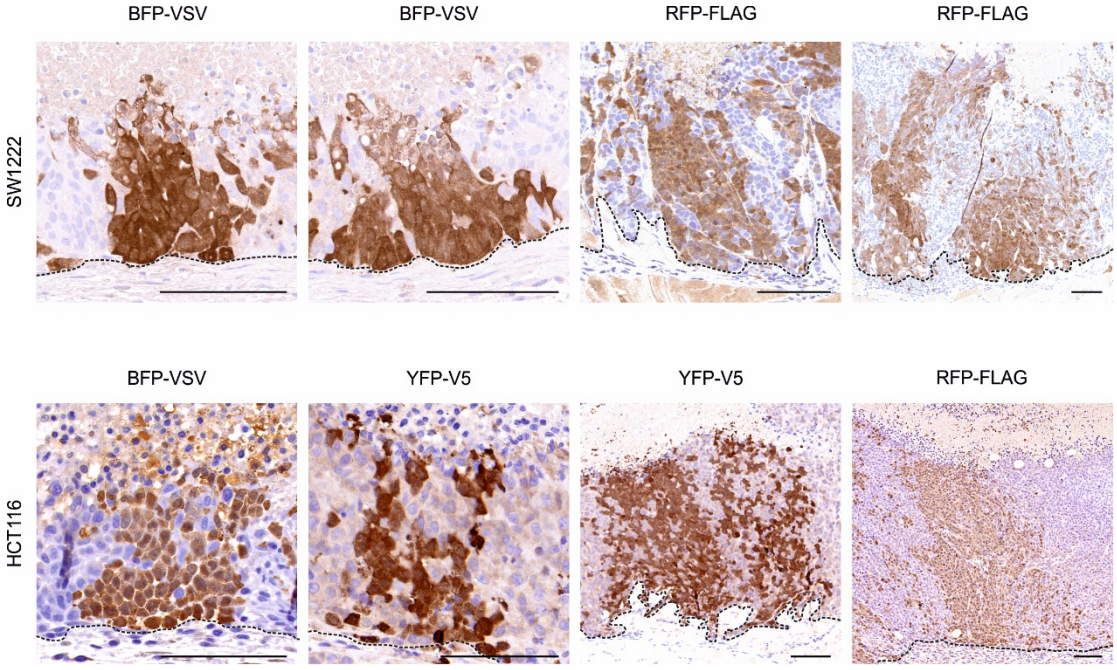

Figure 16: Phenotypes of subclones in colon cancer xenografts. Immunohistochemistry for indicated tagged fluorescence proteins illustrates representative clonal patches in SW1222 and HCT116 xenograft tumors 31 days after recombination. Micrographs show tumors from leading tumor edge (image bottom) to tumor center or central tumor necrosis (image top). Scale bars, 100 µm.

5.1.3. Clone characteristics in colon cancer

To characterize the shape and architecture of colon cancer subclones in more detail, we determined the coordinates of coherent tumor cells with identical colors relative to perpendicular linear axes from the tumor edge to the necrotic tumor center (Figure 17A). An adapted model for linear regression analyses revealed that 31 days after recombination most clones had expanded in a linear manner in SW1222 and HCT116 xenografts, while we observed this with less significance at earlier time points (Figure 17B). Moreover, when we determined the angles (α) of lines fitted to clones by linear regression, relative to tangents to the leading tumor edge (Figure 17A), these predominantly centered around 90° in tumors of both cell lines (Figure 17C).

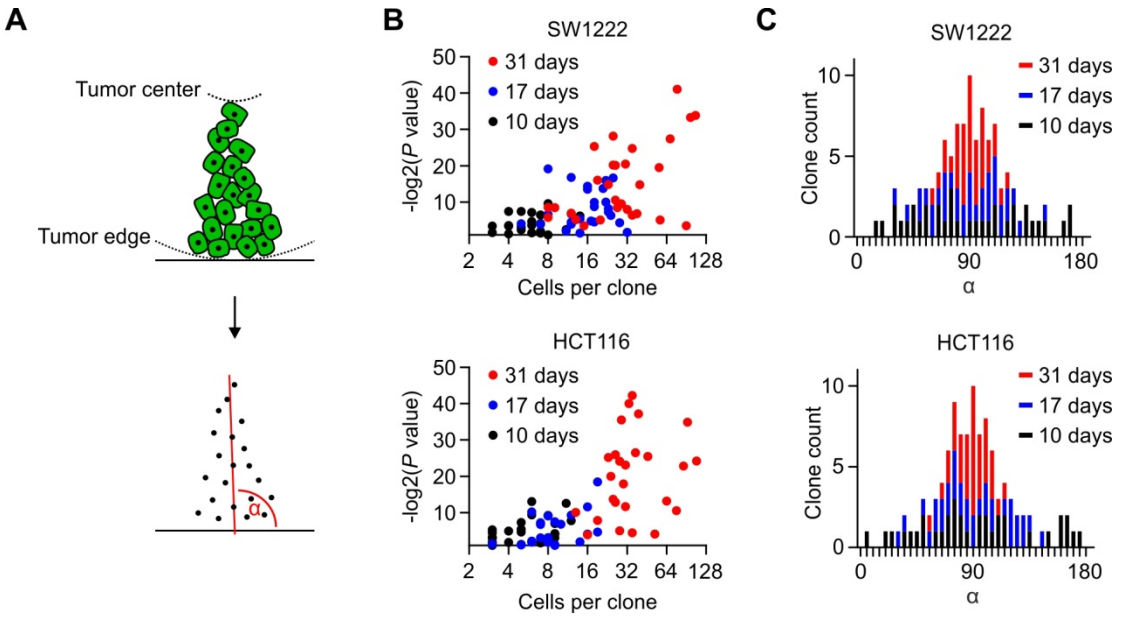

Figure 17: Shape and axis formation of colon cancer subclones. (A) Schematic illustration of clonal analysis. Positions of cells in individual clones relative to leading tumor edge and central tumor necrosis were determined. For each clone, a line of best fit for cell positions then was calculated by linear regression, yielding a clonal axis. α indicates the angle of the clonal axis relative to the leading tumor edge. **(B)** Significance of linear alignment of cells in individual clones at indicated time points after multicolor labeling. **(C)** Angles (α) of clonal axes relative to the leading tumor edge at indicated time points after multicolor labeling

In addition, we then performed BrdU tracing experiments and found that within six or seven days after a single BrdU pulse, the label progressed from the tumor edge toward the tumor center (Figure 18).

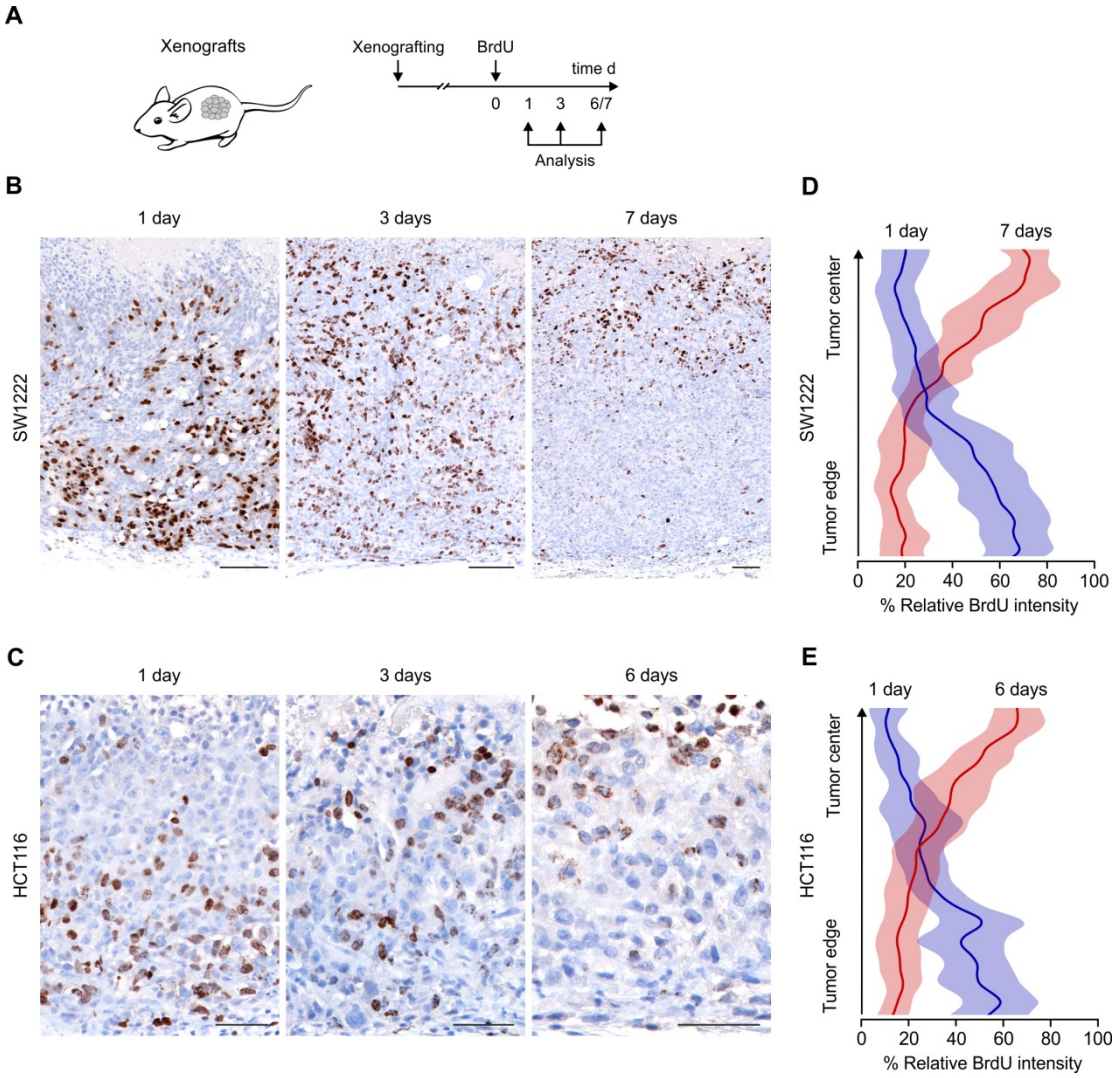

Figure 18: BrdU tracing in colon cancer xenografts. (A) Schema and experimental schedule. **(B, C)** Immunohistochemistry for BrdU in SW1222 and HCT116 xenograft tumors at indicated time points after BrdU pulse. Micrographs show tumors from leading tumor edge (image bottom) to tumor center or central tumor necrosis (image top). Scale bars, 100 µm in **(B)** and 50 µm in **(C)**. **(D, E)** Quantification of average relative BrdU staining intensity from tumor edge to tumor center. Data are mean with 95 % confidence bands and derived from different xenograft areas (n ≥ 20) of independent biological replicates (n ≥ 3).

Collectively, these findings provided evidence of a non-random linear expansion of tumor cell clones, perpendicular to the leading tumor edge and directed towards the tumor center.

With these findings in mind, we analyzed the distribution of nuclear β-catenin, FRA1, CK20, and GLUT1 within individual clones. As expected for organized SW1222-derived xenografts, nuclear β-catenin and FRA1 marked tumor cells predominantly at the leading tumor edge, while CK20 and GLUT1 marked cells close to the

necrotic tumor center within these clones, indicating clonal axis formation along the centripetal differentiation axis in these tumors (Figures 19 and 20).

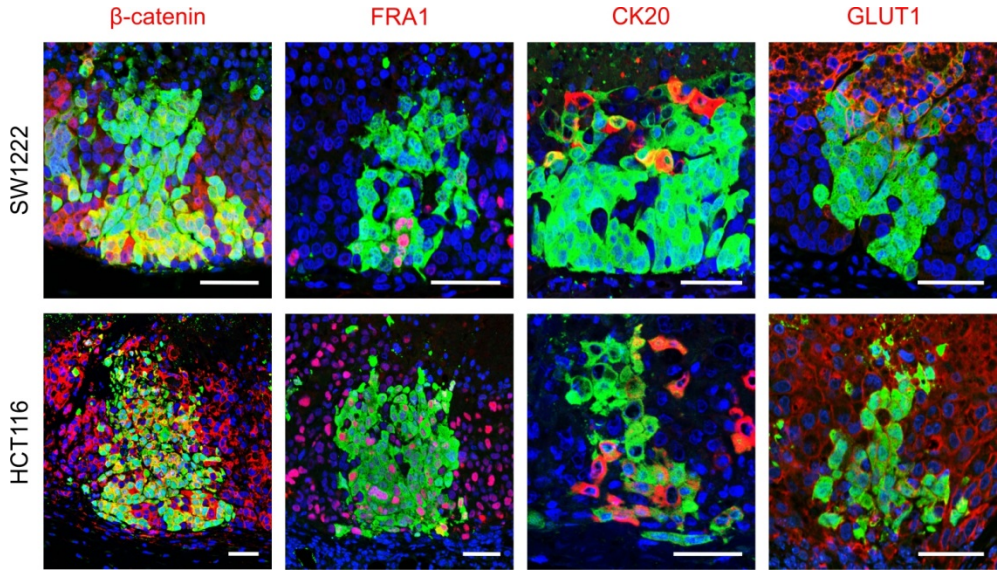

Figure 19: Marker distribution in CRC Xenografts. Confocal images show positions of indicated stem cell and differentiation antigens (red) in individual clones (green) of colon cancer xenografts. Fluorescent images show xenograft tumors from leading tumor edge (image bottom) to central tumor necrosis (image top). Scale bars, 50 µm.

Importantly, however, since clonal axes in disorganized HCT116 colon cancer xenografts also were perpendicular to the leading tumor edge and all four markers were more randomly expressed in individual clones of these tumors, this indicated that clonal axis formation does not generally parallel or depend on differentiation gradients. These findings suggested limited influence of differentiation gradients on clonal architecture and outgrowth in colon cancer. (Figures 19 and 20).

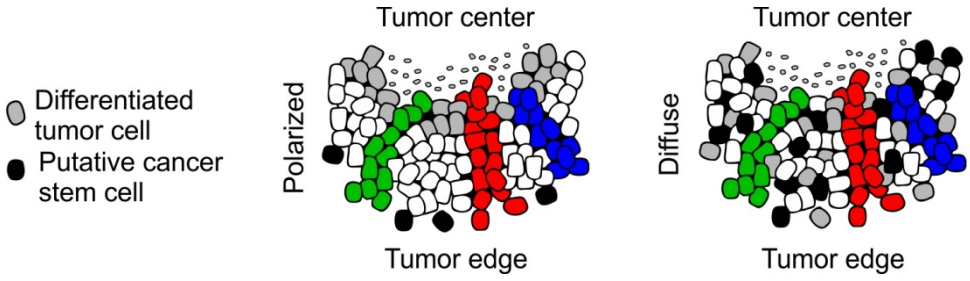

Figure 20: Differentiation gradients in colon cancers. Schematic model suggesting identical clonal outgrowth in organized and disorganized colon cancers with and without centripetal differentiation gradients, respectively.

5.1.4. Clonal dynamics in colon cancer

To further learn about clonal dynamics in colon cancer xenografts, we analyzed clone sizes and clonal density after multicolor labeling over time. Three days after recombination, clones were composed of two-three cells in average. Clone sizes then increased exponentially until 17 days with subsequently slightly slowed growth rates (Figure 21). Accordingly, clonal density, i.e. the number of clones per area, decreased over time. Importantly, when comparing clonal density at the tumor edge and close to the central tumor necrosis, we observed a significantly earlier decrease in clonal density at the leading tumor edge, most obvious at 10 days and 17 days after recombination in both SW1222 and HCT116 colon cancer xenografts (Figure 21).

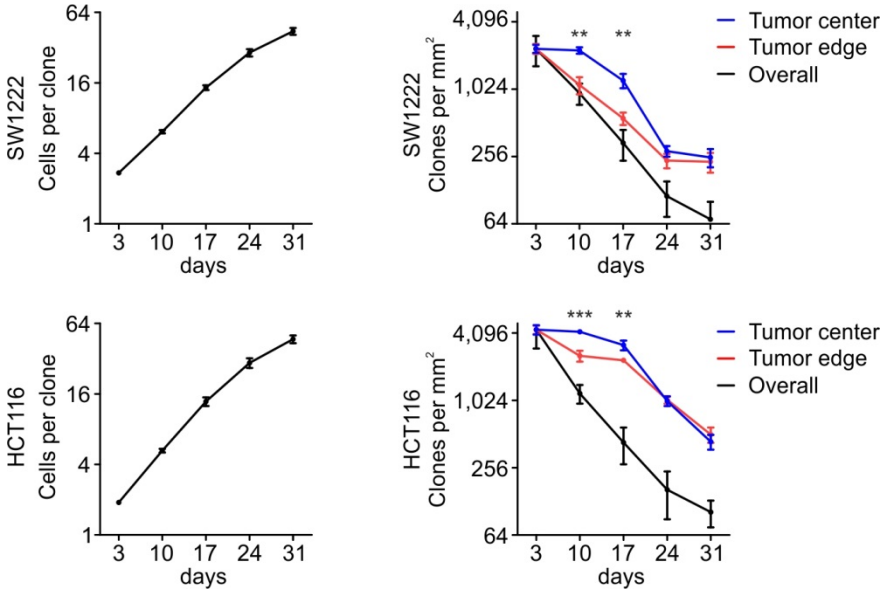

Figure 21: Clonal analysis *in vitro* over time. Clone sizes (left panels) and clones per area (right panels), as determined by analysis of confocal fluorescence images of SW1222 and HCT116 colon cancer xenografts at different time points after multicolor labeling. Clones per area were measured overall and separately in tumor thirds close to the tumor edge and close to central tumor necrosis, as indicated. *** P < 0.001 and ** P < 0.01 indicate differences between tumor edge and tumor center by t-test.

Also, over time the average clonal width at the leading tumor edge linearly increased when adjusted to increases in tumor circumference (Figure 22).

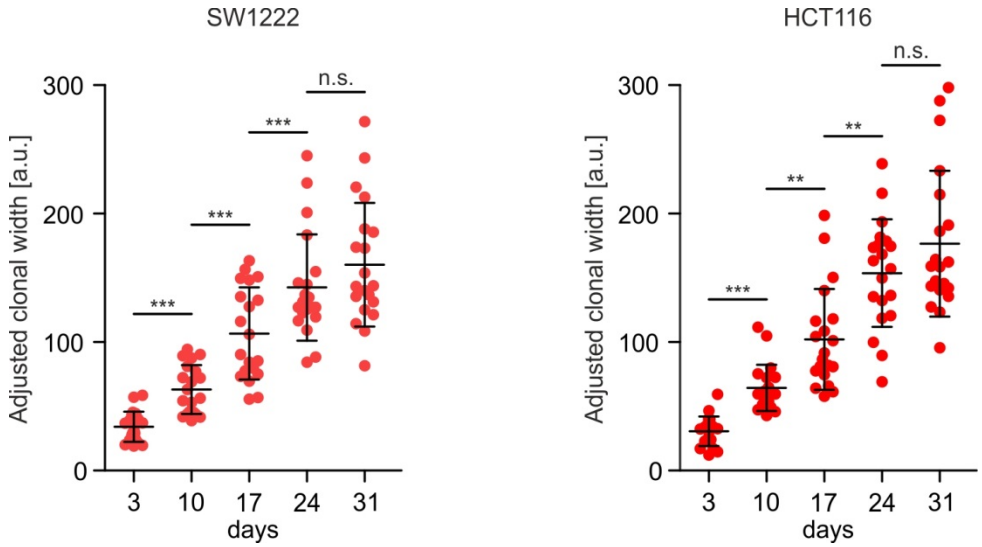

Figure 22: Lateral clonal expansion in colon cancer xenografts. Measurements of clonal width at the leading tumor edge divided by relative changes in tumor circumference (Adjusted clonal width) are shown for indicated time points after multicolor recombination. a.u, arbitrary units. Error bars indicate mean ± SD. **, P < 0.01; ***, P < 0.001; n.s., nonsignificant by t test.

Together with the observation that clones could be lost into central tumor necrosis, these findings suggested that clonal competition mainly occurred at the leading tumor edge with subsequent clonal outgrowth towards the necrotic tumor center. Of note, when we analyzed individual clones 31 days after recombination for cancer hot spot mutations, no mutational differences were observed. This indicated that clonal outgrowth and competition likely occured in the absence of overt changes in driver mutation profiles (Table 1).

Table 1: Mutational status of individual microdissected subclones of SW1222 and HCT116 xenograft tumors, 31 days after recombination

		SW1222 clones					HCT116 clones				
Gene	**Mutation**	**#1**	**#2**	**#3**	**#4**	**#5**	**#1**	**#2**	**#3**	**#4**	**#5**
KRAS	G13D	x	x	x	x	x					
CTNNB1	S45 del	x	x	x	x	x					
PIK3CA	H1047R	x	x	x	x	x					
SMO	V404M	x	x	x	x	x					
ABL1	D276G	x	x	x	x	x					
KRAS	A146V						x	x	x	x	x
APC	G1306T						x	x	x	x	x

Finally, we inferred a two-dimensional spatial simulation model for clonal dynamics in colon cancer, implementing few rules only that we derived from our *in vivo* observations (Figure 23).

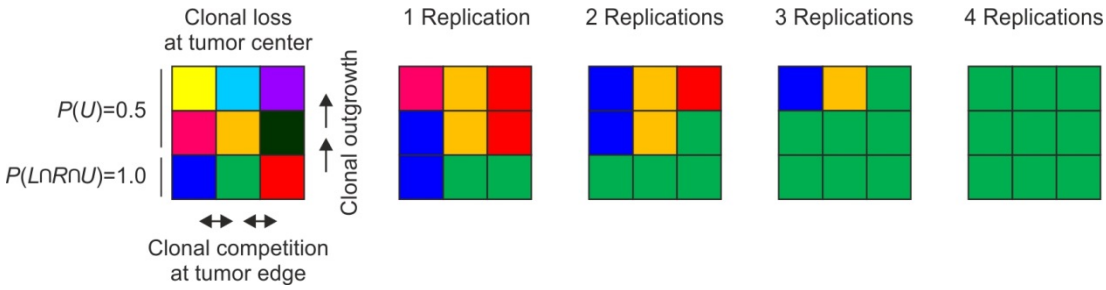

Figure 23: Two-dimensional simulation model for clonal outgrowth. Bottom row simulates tumor cells at the tumor edge and top row simulates tumor cells neighboring tumor necrosis. At the tumor edge cells divide during each simulated replication cycle and probabilities for cell expansion upwards P(U), to the left P(L) or to the right P(R) are equal. In other positions, cells divide upwards only at decreased frequency with P(U) = 0.5. Panels simulate one possible outcome for four replication cycles.

First, clones may only be lost into the tumor center or into central tumor necrosis, represented by the upper border of our square model. Second, clonal competition by lateral clone expansion may only occur at the leading tumor edge, represented by the lower border of the model. Third, based on measurements of proliferation by Ki67 in primary colon cancers (n=92) and xenograft tumors (Figure 24), growth rates in our model were slowed to 0.5 in central tumor areas relative to the leading tumor edge.

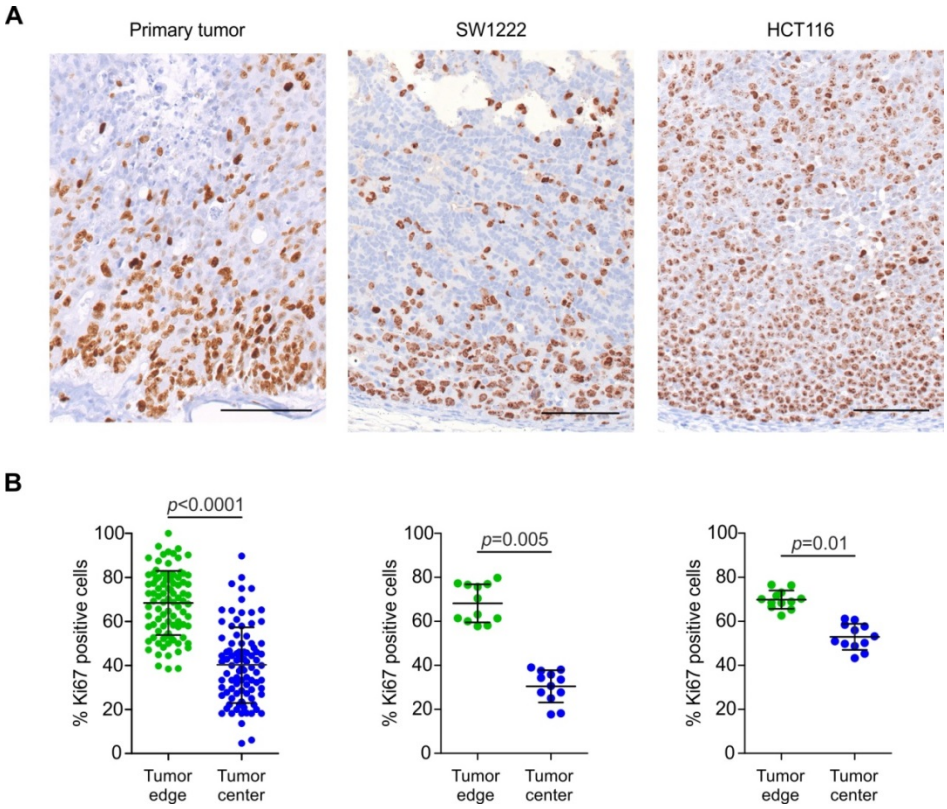

Figure 24: Proliferation gradients in colon cancer. (A) Representative immunohistochemistry and **(B)** quantification of Ki67 in primary colon cancers (n = 92) and in SW1222 and HCT116 xenograft tumors (4 different areas in 3 biological replicates), as indicated. Micrographs show tumors from leading tumor edge (image bottom) to tumor center or central tumor necrosis (image top). Scale bars, 100 µm. Error bars indicate mean ± SD. P values are t test results.

This model, when composed of few “cells” only, illustrated rapid loss of individual tumor cells and a drift towards mono-clonality (Figure 25).

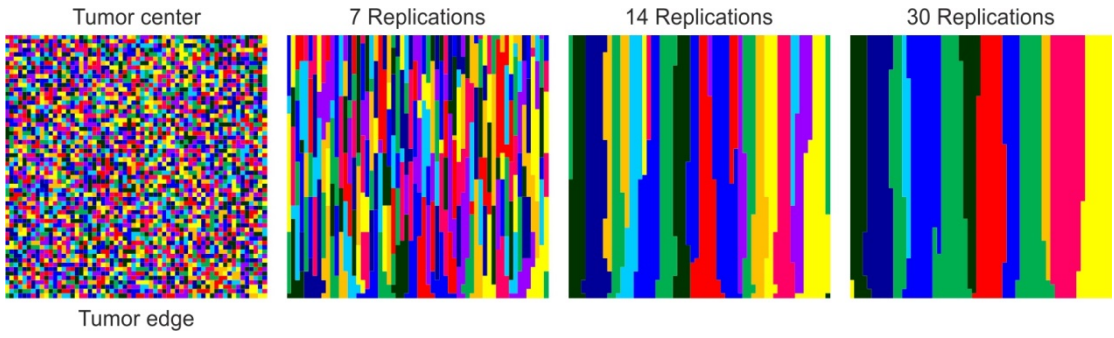

Figure 25: Two-dimensional simulation model for clonal outgrowth. Bottom row simulates tumor cells at the tumor edge and top row simulates tumor cells neighboring tumor necrosis. At the tumor edge cells divide during each simulated replication cycle and probabilities for cell expansion upwards P(U), to the left P(L) or to the right P(R) are equal. In other positions, cells divide upwards at decreased frequency with P(U) = 0.5. One possible outcome after 30 replication cycles is illustrated.

In larger scale, linear expansion of tumor cell clones from the leading tumor edge towards the central tumor necrosis were seen with widening of some clones and inevitable loss of those that lost contact to the leading edge, causing a continuous drift towards oligo-clonality, well-fitting our *in vivo* findings (Figure 25).

Also, the dynamics of gains in clone size and loss in clone density over time quantitatively matched our observations in colon cancer xenografts (Figure 26).

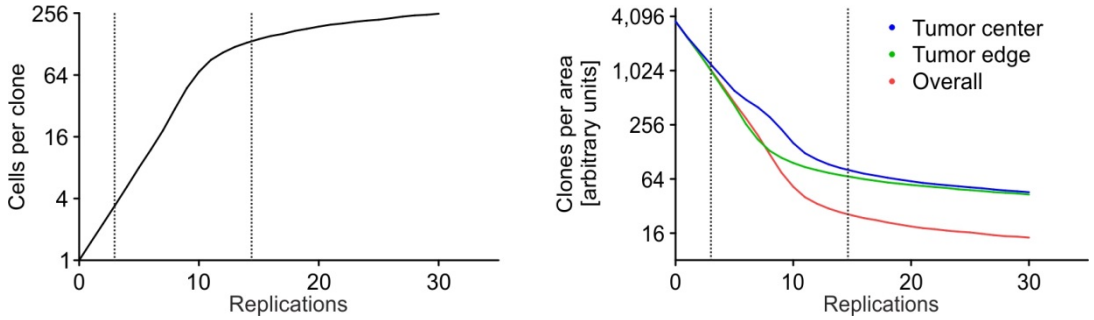

Figure 26: Clonal analysis within the simulation model. Average clone sizes (left panel) and clones per area (right panel) from 100 independent simulations. Clones per area (arbitrary units) are given overall, and in thirds of the model close to tumor edge and central tumor necrotic core, respectively. Dotted lines approximately deliminate simulation segments fitting our *in vivo* data.

Importantly, this also included an earlier decrease of clonal density at the leading tumor edge compared to the tumor center, which was due to the implemented restriction of clonal competition to the leading tumor edge. Collectively, this model corroborated the idea that the *in vivo* observed clonal outgrowth from the leading tumor edge towards the tumor center may be based on few rather positional characteristics of colon cancer cells while differentiation gradients may be of less importance.

5.2. Identification of PBX3 as an inducer of EMT signaling in colon cancer

The results presented in this section are part of the publication: **Lamprecht S**, Kaller M, Schmidt EM, Blaj C, Hermeking H, Grünewald T G, Jung A, Kirchner T, Horst D. PBX3 is part of an EMT regulatory network in colorectal cancer and indicates poor outcome. Clin Cancer Res. 2018; 24(8):1974-1986

5.2.1. PBX3 is overexpressed in colon cancer cells with high WNT activity

To find transcription factors linked to WNT signaling activity in colon cancer, we screened three previously published gene expression data sets that were derived from colon cancer cell subpopulations with low and high WNT activity (Horst et al., 2012; Vermeulen et al., 2010). Of 956 represented genes that encoded for known or putative transcription factors, 69 (7.2 %) were significantly ($P < 0.05$ by t-test) differentially expressed by 1.5 fold or more. Among those with most consistent overexpression in tumor cells with high WNT activity that expectedly included known WNT pathway components or target genes such as LEF1, PROX1, and ZEB1, we identified PBX3 (Figure 27, Table 2).

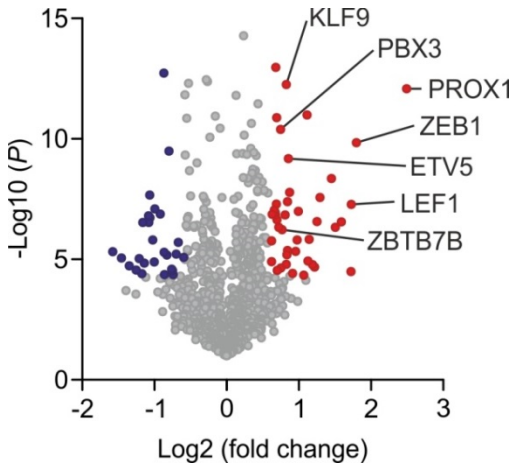

Figure 27: Volcano plot of protein expression. Volcano plot of gene expression data of genes encoding for transcription factors, derived from three pooled data sets comparing colon cancer cells with high and low WNT activity. Colored dots denote genes that in each data set are significantly ($P < 0.05$) upregulated (red) or downregulated (blue).

Table 2: Differentially expressed genes (F.C.≥1.5; P≤0.05) encoding for known or putative transcription factors derived from three gene expression data sets of colon cancer cells with high vs. low WNT activity. P values are t test results.

Gene Symbol	F.C. WNT high vs. low	*P* value WNT high vs. low
PROX1	5.64	0.0002
ZEB1	3.48	0.0011
LEF1	3.31	0.0064
DLX5	3.3	0.0443
TCF4	3.01	0.0106
HOXD9	2.84	0.0124
IRX5	2.74	0.003
NFE2	2.45	0.0053
ETV1	2.38	0.0105
ELF5	2.33	0.0389
SNAI2	2.28	0.0363
NKX2-1	2.21	0.0176
HEY2	2.19	0.0327
MEIS1	2.16	0.0005
SIX2	2.1	0.0494
HNF1B	1.99	0.0079
GBX2	1.97	0.0179
ZIC4	1.94	0.0248
PHOX2B	1.88	0.0467
TBX10	1.83	0.0046
ETV5	1.81	0.0017
EGR3	1.8	0.006
TEF	1.8	0.0242
EGR4	1.79	0.0272
KLF9	1.77	0.0002
ZNF467	1.77	0.0359
SMARCA1	1.75	0.0087
ZBTB7B	1.7	0.0133
ZNF287	1.69	0.04
PBX3	1.68	0.0007
ZFP30	1.65	0.0125
ZFP2	1.63	0.0427
NFATC1	1.62	0.01
ELK3	1.62	0.0005
ZNF236	1.61	0.0064
KLF12	1.6	0.0001
NKX3-1	1.6	0.0087
ZBTB1	1.6	0.008
ZBTB20	1.58	0.0076
ZBTB44	1.55	0.0086
PAX8	1.54	0.0335
DNMT3B	1.54	0.0183
MIER2	0.66	0.0296
NFKBIL1	0.63	0.0192
RARB	0.62	0.0269
FLI1	0.6	0.0485
FOXE1	0.59	0.0431
SIX5	0.59	0.0414
NR3C2	0.57	0.0014

Gene Symbol	F.C. WNT high vs. low	*P* value WNT high vs. low
ZNF154	0.55	0.0484
HOXB1	0.55	0.0254
BHLHE41	0.55	0.0001
NR1H4	0.53	0.0085
MYT1	0.5	0.0073
MEF2C	0.5	0.0337
SOX17	0.49	0.0178
TFEC	0.48	0.0094
CREB3L1	0.48	0.0049
ESR1	0.47	0.0109
NKX6-1	0.47	0.0089
GATA4	0.45	0.0347
NEUROG1	0.45	0.0108
FEV	0.44	0.0471
MAF	0.43	0.0307
ZXDB	0.42	0.0424
SPDEF	0.39	0.0377
FOXG1	0.36	0.03
MEOX2	0.34	0.0251

Indeed, in this data set increased PBX3 expression coincided with high expression of WNT pathway components and target genes and, conversely, with repression of genes associated with a differentiated tumor cell phenotype (Figure 28).

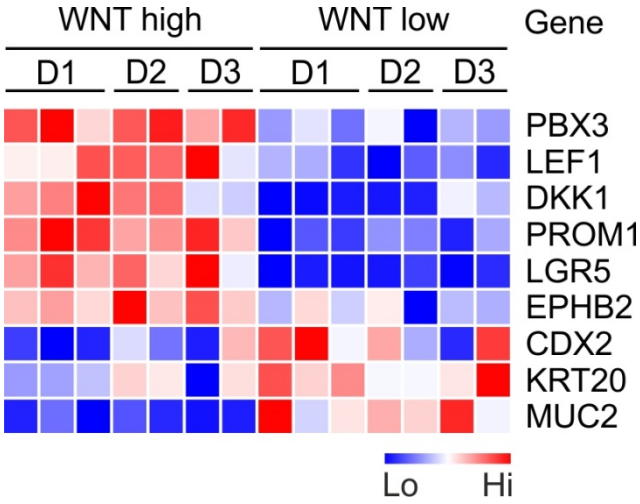

Figure 28: PBX3 is overexpressed in colon cancer cells with high WNT activity. Heatmaps of PBX3, selected WNT targets, and differentiation factors in three data sets (D1–D3) of colon cancer cells with high and low WNT activity.

We then examined tissues sections of individual colon cancers and found that PBX3 was heterogeneously expressed with strongest expression at the leading tumor edge, where it overlapped with strong expression of nuclear β-catenin, indicating high WNT activity. These findings identified upregulation of PBX3 in colorectal cancer cells with high WNT activity on mRNA and protein levels, and suggested a possible regulation of PBX3 by WNT (Figure 29).

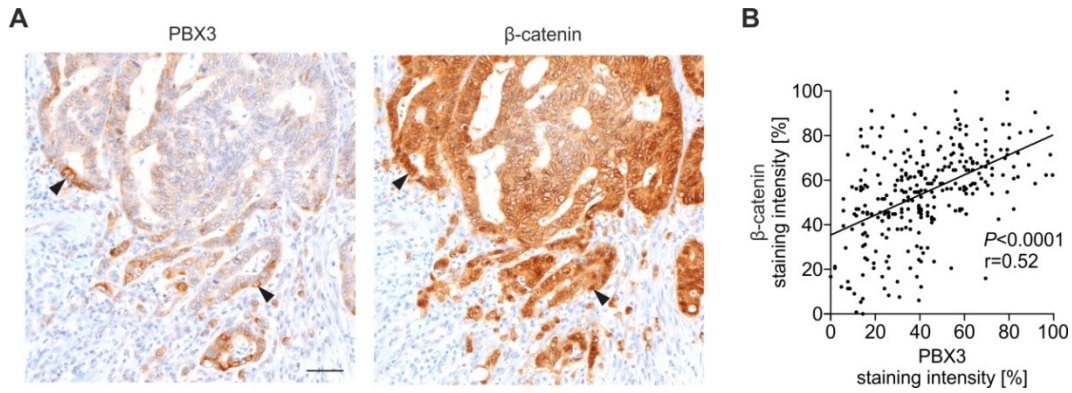

Figure 29: Correlation of PBX3 and β-catenin. (A) Immunohistochemical staining of serial sections illustrate upregulation of PBX3 in areas with increased β-catenin staining (arrows). scale bar, 50 µm. **(B)** Quantification of immunohistochemical signals for PBX3 and β-catenin. Values are given as staining intensity for individual tumor cells (n ≥ 300) of different colorectal cancer samples (n ≥ 6). *P* values are results of linear regression analyses.

5.2.2. PBX3 expression is regulated by WNT signaling in colorectal cancer

Next, we determined whether PBX3 expression depended on high WNT activity in colon cancer. Reducing WNT activity by depletion of β-catenin with two different siRNAs reduced PBX3 in different colon cancer cell lines on the protein level, whereas transfection with control siRNA had no effect on PBX3 (Figure 30).

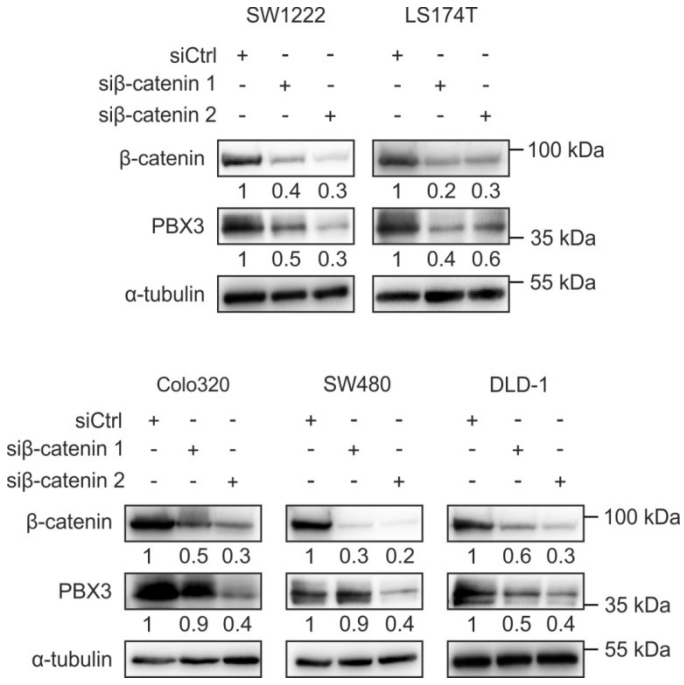

Figure 30: β-catenin knockdown reduced PBX3 expression. Immunoblotting of indicated proteins after transfection of different colon cancer cells with siRNA against β-catenin. Numbers below immunoblots indicate fold change by densitometry.

We confirmed these effects on the mRNA level by qRT-PCR in two cell lines, in which β-catenin knockdown significantly downregulated *PBX3* expression and that of the WNT target genes *AXIN2*, *NKD1* and *LGR5* (Figure 31) (Barker et al., 2007; Yan et al., 2001).

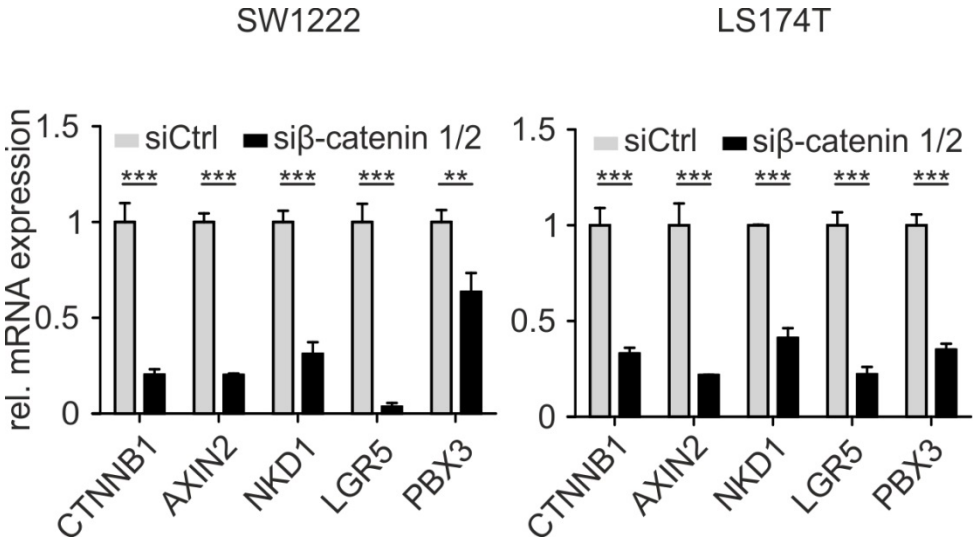

Figure 31: β-catenin knockdown represses PBX3 expression. qRT-PCR results on indicated genes after knockdown of β-catenin. Error bars indicate mean ± SD. **, P < 0.01; ***, P < 0.001 by t test, n ≥ 3.

Furthermore, we tested the effects of a doxycycline (DOX) inducible dominant negative TCF4 (dnTCF4), a potent inhibitor of the β-catenin/TCF4 transcription factor complex (Van de Wetering et al., 2002). In two cell lines, dnTCF4 induction strongly reduced transcription from β-catenin/TCF4 binding sites, as seen in TOPflash luciferase reporter assays (Figure 32, left panel) but also decreased PBX3 protein expression and downregulated *PBX3* mRNA levels among the panel of WNT target genes (Figure 32, middle and left panel).

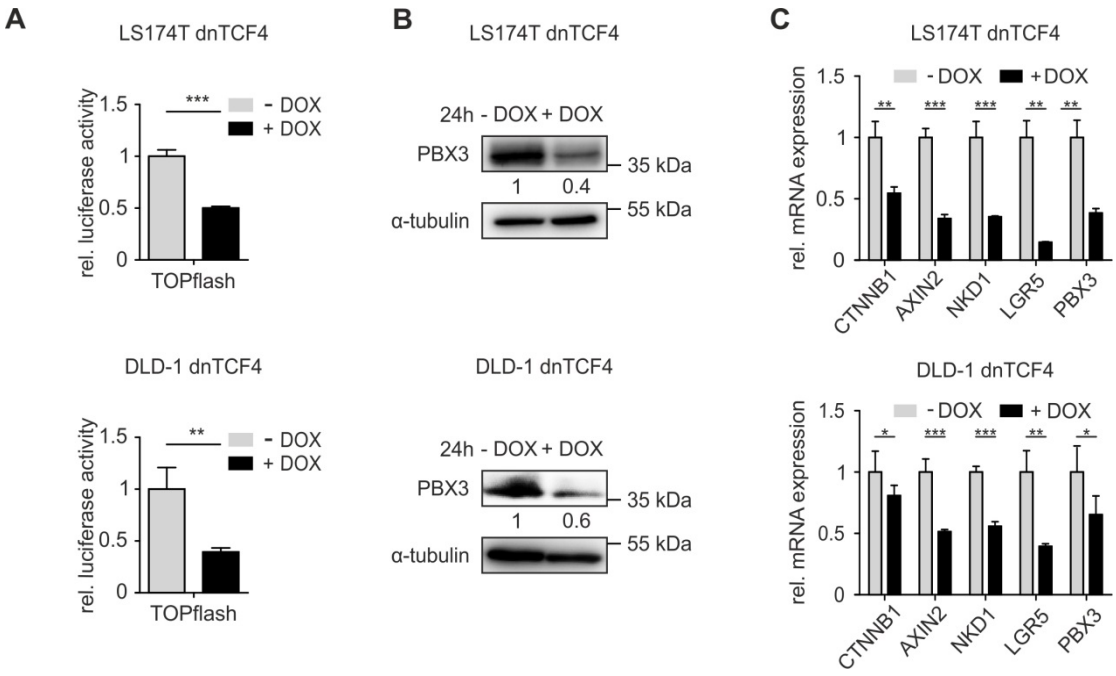

Figure 32: Effects of WNT pathway inhibition on PBX3. Immunoblotting **(A)**, Dual-luciferase assays with TOPflash reporter constructs **(B)** and qRT-PCR **(C)** results on indicated proteins or genes after inhibition of the β-catenin/TCF4 by doxycycline dependent induction dnTCF4 in HEK293 cells. Error bars indicate mean ± SD. *, P < 0.05; **, P < 0.01; ***, P < 0.001 by t test, n ≥ 3.

On the contrary, stimulation of HEK293T, a cell line with low intrinsic WNT activity, with WNT3a led to strong overexpression of PBX3 and active β-catenin on the protein level, as well as upregulation of *PBX3* and WNT target gene mRNA (Figure 33).

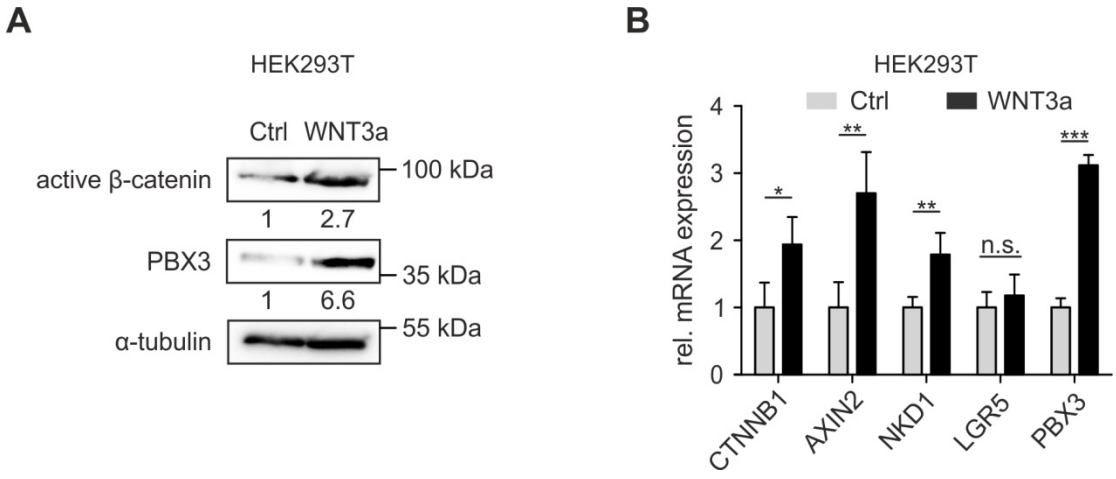

Figure 33: Effects of WNT pathway activation on PBX3. Immunoblotting **(A)** and qRT-PCR **(B)** results on indicated proteins or genes after stimulation of HEK293 cells with WNT3a. Numbers below immunoblots indicate fold change by densitometry. Error bars indicate mean ± SD. *, P < 0.05; **, P < 0.01; ***, P < 0.001; n.s., not significant by t test , n ≥ 3.

Taken together these findings suggested that PBX3 expression is regulated by WNT signaling in colon cancer, while this effect is not cell type specific.

To test for direct transcriptional regulation of PBX3 by WNT, we then screened the *PBX3* promoter sequence and identified three putative β-catenin/TCF4 binding motifs (WWCAAAG (Korinek et al., 1998)) within 2.5 kb 5' of the first exon of the PBX3 gene (Figure 34).

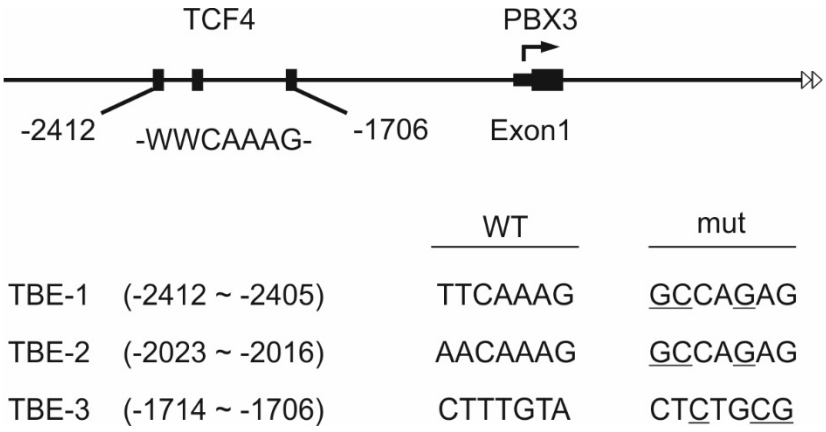

Figure 34: TCF4 binding motifs within the PBX3 promoter sequence. three putative β-catenin/TCF4 binding motifs were found within within 2.5 kb of the PBX3 promoter sequence.

To determine if PBX3 is induced by WNT activation via these β-catenin/TCF4 motifs, we subjected 2.5 kb of the *PBX3* promoter region including these motifs, or mutated motifs as control, to dual luciferase reporter assays. Unexpectedly, WNT3a stimulation did not increase luciferase expression from the wild-type reporter (Figure 35, A), while TOPflash assays confirmed strong transcriptional activation of WNT signaling by WNT3a (Figure 35, B). These data suggested that *PBX3* is no direct β-catenin/TCF4 target gene but instead modulated by other WNT dependent downstream factors.

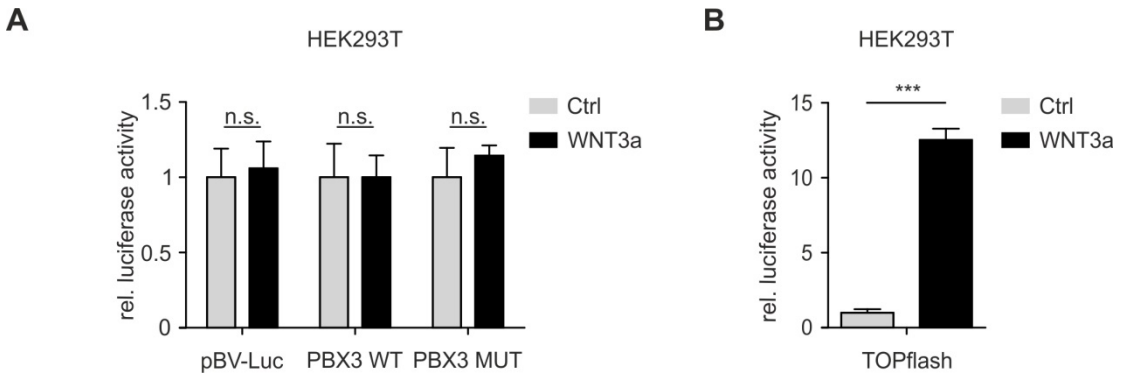

Figure 35: No effects of WNT activation on β-catenin/TCF4 motifs within the PBX3 promoter. Dual-luciferase assays with wildtype or mutated reporter constructs of the PBX3 promoter **(A)** and TOPflash reporter constructs **(B)** after stimulation of HEK293 cells with WNT3a. Error bars indicate mean ± SD. ***, P < 0.001; n.s., not significant by t test, n ≥ 3.

5.2.3. PBX3 is strongly associated with EMT in colon cancer

Since PBX3 expression was strongest in tumor cells located at the infiltrative tumor edge with high WNT activity (Figure 29), and WNT signaling is a known regulator of EMT in colon cancer (Brabletz et al., 2005), we hypothesized that PBX3 might be linked to EMT. To test for a general association of PBX3 with an EMT phenotype, we assembled and normalized publicly available mRNA expression data of 1.097 colon cancers. Supporting our hypothesis, Gene Set Enrichment Analyses (GSEA) conceded highly significant (p<0.001) overlap of *PBX3* expression and the expression of two published core EMT gene signatures (Figure 36) (Anastassiou et al., 2011; Taube et al., 2010).

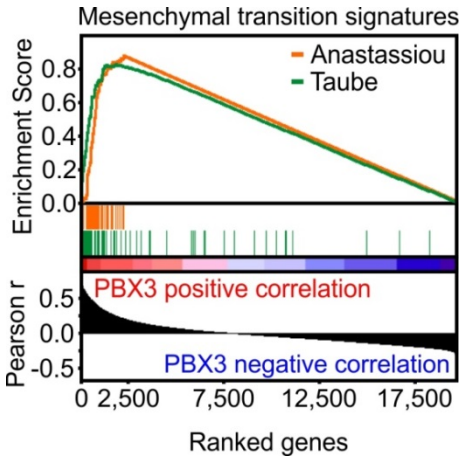

Figure 36: PBX3 correlates with a mesenchymal transition signature. GSEA for genes ranked by Pearson correlation (Pearson r) to *PBX3* expression for two EMT target gene signatures by Anastassiou and colleagues (orange curve: ES = 0.89, P < 0.001) and Taube and colleagues (green curve: NES = 0.84, P = 0.001) in 1.097 RNA-Seq datasets of colon cancer from TCGA.

Additionally, factors that reportedly imply EMT activity in CRC were upregulated in tumors that showed high levels of *PBX3* expression, among them most prominently *ZEB1* (r=0.69, p<0.0001), a known inducer of EMT in colon cancer cells (Figure 37).

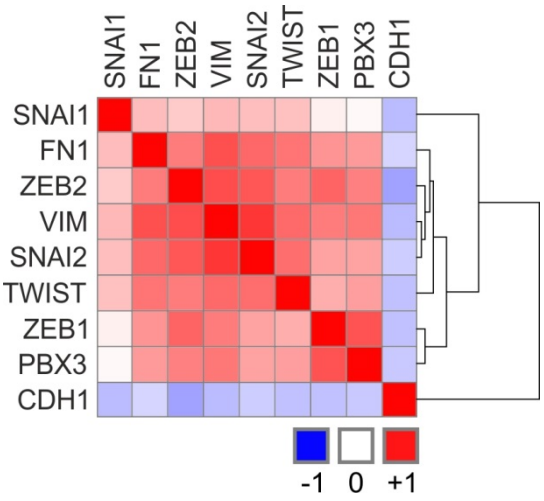

Figure 37: PBX3 is overexpressed in colon cancers with high expression of EMT related genes. Heatmaps of *PBX3*, selected mesenchymal and epithelial targets in a dataset of 1.097 colon cancers. Colors represent Pearson r from -1 (blue) to 1 (red).

On the contrary, *CDH1,* the epithelial differentiation marker negatively correlated with *PBX3*, further supporting the idea of an association of *PBX3* and EMT. To shed more light on PBX3 and EMT *in situ*, we subsequently assessed colon cancer tissues for PBX3 and LAMC2, a factor regulated by ZEB1 (Sanchez-Tillo et al., 2011) by immunohistochemistry, and found a highly significant correlation of both markers (Figure 38). Taken together these data demonstrated that PBX3 is associated with an EMT tumor cell phenotype in colon cancer (Peinado et al., 2007; Sánchez-Tilló et al., 2012).

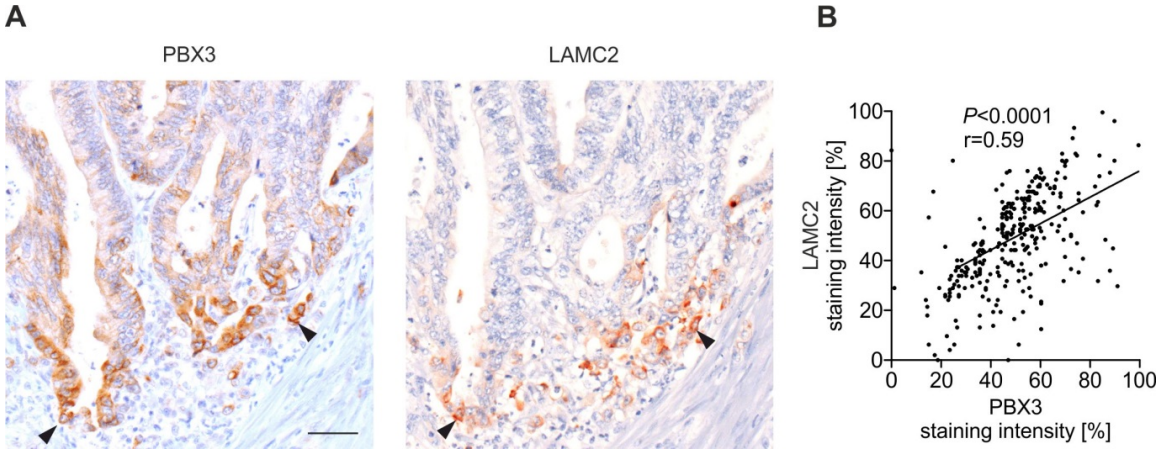

Figure 38: Correlation of PBX3 and LAMC2. (A) Immunohistochemical staining of serial sections illustrate upregulation of PBX3 in areas with increased LAMC2 staining (arrows). scale bar, 50 µm. **(B)** Quantification of immunohistochemical signals for PBX3 and LAMC2. Values are given as

staining intensity for individual tumor cells (n ≥ 500) of different colorectal cancer samples (n ≥ 5). P values are results of linear regression analyses

5.2.4. PBX3 is induced by EMT in colon cancer and required for a full EMT phenotype

EMT can be induced in colon cancer cells by ectopic expression of SNAIL or ZEB1 (Peinado et al., 2007), and we applied this approach to test if PBX3 expression is EMT dependent. We used a DOX inducible episomal vector system to overexpress either SNAIL or ZEB1 in DLD-1 and LS174T cells, two colon cancer cell lines with low EMT marker expression and pronounced epithelial phenotypes (Jackstadt et al., 2014). In both cell lines, induction of SNAIL by DOX treatment caused upregulation of PBX3 protein levels within 12 h and also increased *PBX3* mRNA levels together with *VIM* and *ZEB1*, while repressing *CDH1* and *miR-200c*, indicating an EMT phenotype (Figure 39).

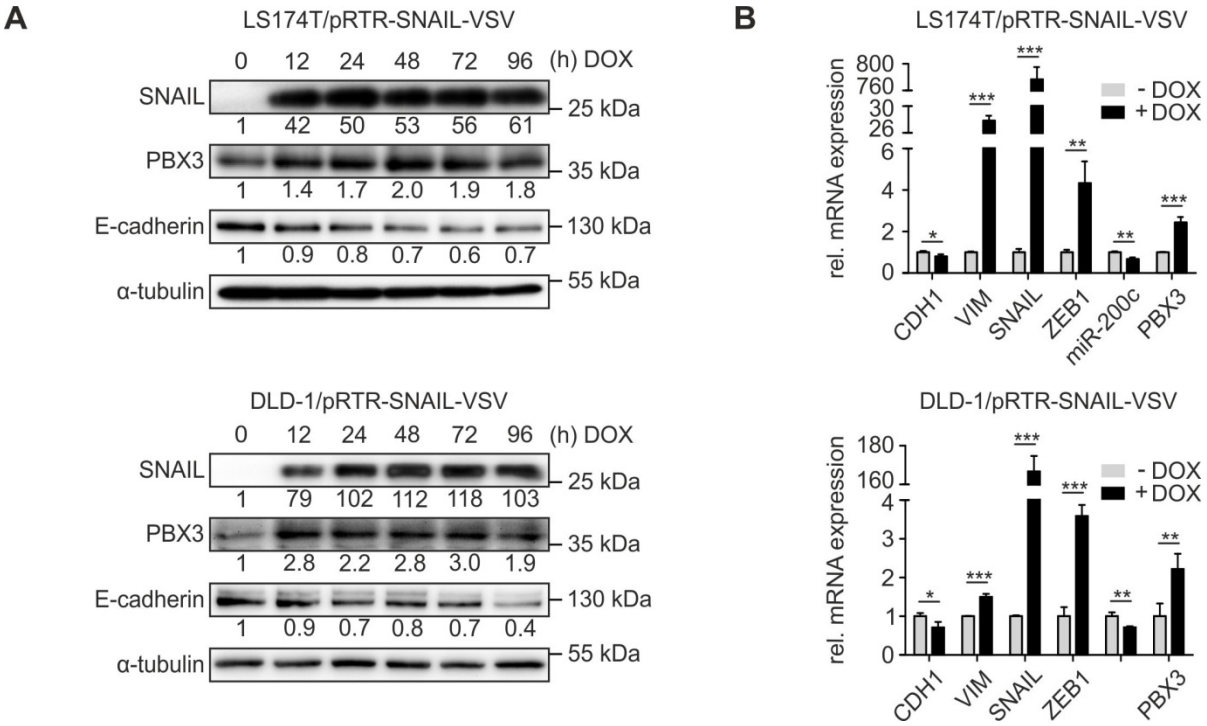

Figure 39: SNAIL induces PBX3 expression. (A) Western blot analysis of PBX3 protein levels in LS174T and DLD-1 cells with a pRTR-SNAIL-VSV vector after treatment with DOX for the respective periods. Numbers below immunoblots indicate fold change by densitometry. **(B)** Gene expression analyses by qRT-PCR for indicated genes after 72 h DOX induction. Error bars represent mean ± SD. *, P < 0.05; **, P < 0.01; ***, P < 0.001 by t test, n ≥ 3.

Immune fluorescence further confirmed upregulation of PBX3 together with rearranged F-actin stress fibers, a characteristic feature of the EMT process (Figure 40) (Moreno-Bueno et al., 2009).

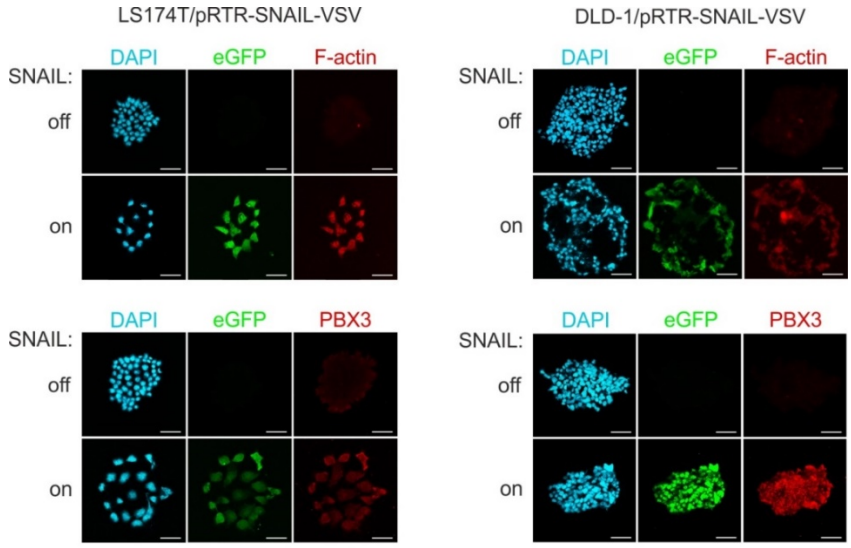

Figure 40: Clonal analysis within the simulation model. Representative confocal immunofluorescence images of LS174T and DLD-1 cells with a pRTR-SNAIL vector after treatment with DOX for respective proteins and DAPI as nuclear counterstain. Scale bars, 50 µm.

Similarly, ZEB1 induction also caused upregulation of PBX3 protein and mRNA levels with downregulation of *miR-200c* in both cell lines, and *CDH1* repression in DLD-1 cells, while it had less effect on the other EMT markers (Figure 41).

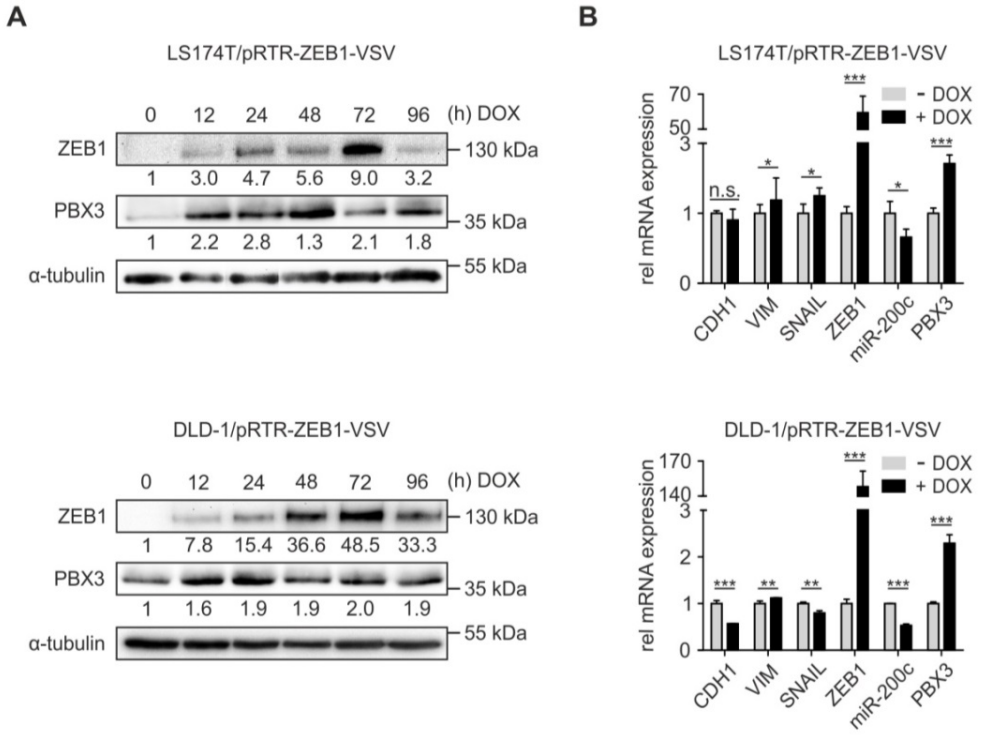

Figure 41: ZEB1 induces PBX3 expression. (A) Western blot analysis of PBX3 protein levels in LS174T and DLD-1 cells with a pRTR-ZEB1-VSV vector after induction with DOX for the respective periods. Numbers below immunoblots indicate fold change by densitometry. **(B)** PBX3 mRNA levels in the same experiment analogous to (A). Numbers below immunoblots indicate fold change by densitometry. Error bars represent mean ± SD. *, P < 0.05; **, P < 0.01; ***, P < 0.001 by t test, n ≥ 3.

Because this suggested a regulation of PBX3 through the SNAIL-ZEB signaling axis (Thiery et al., 2009), we next examined the effects of ZEB1 depletion on PBX3 in SW480 and Colo320 colon cancer cells, both of which have high levels of ZEB1 and a mesenchymal phenotype, when compared to DLD-1 and LS174T (Hahn et al., 2013). Indeed, ZEB1 depletion by siRNA decreased PBX3 protein and mRNA while *miR-200c* significantly increased in both cell lines (Figure 42). PBX3 therefore not only correlates with a mesenchymal phenotype but also is induced by EMT in colon cancer, while this appears to depend on ZEB1.

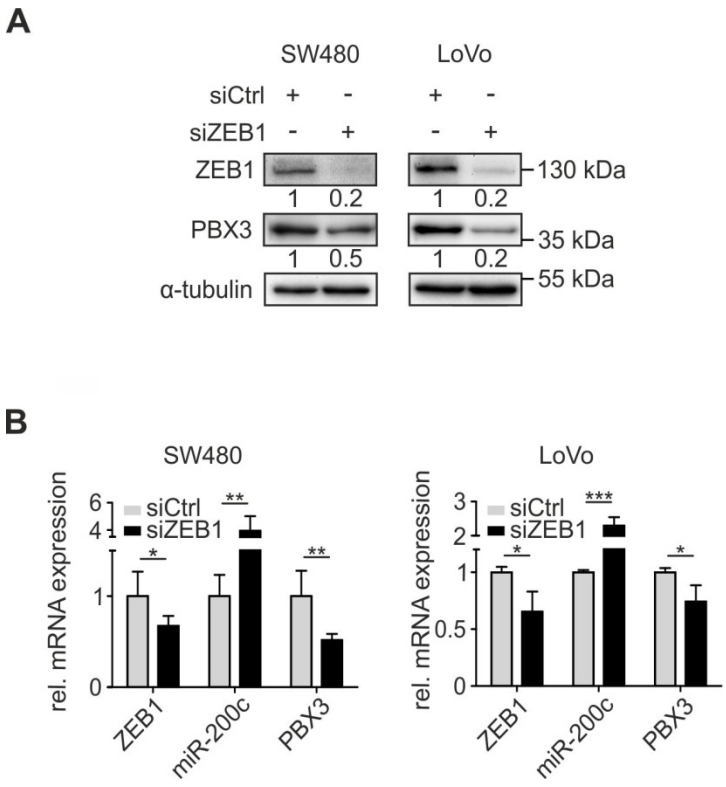

Figure 42: ZEB1 knockdown represses PBX3 expression. (A) Immunoblotting of indicated proteins after transfection of SW480 and LOVO colon cancer cells with siRNA against ZEB1. Numbers below immunoblots indicate fold change by densitometry. **(B)** PBX3 mRNA levels in the same experiment analogous to (A). Error bars represent mean ± SD. *, P < 0.05; **, P < 0.01; ***, P < 0.001 by t test, n ≥ 3.

Because EMT induction by ZEB1 causes repression of *miR-200*, and *PBX3* is a recently identified *miR-200* target, we asked whether the effects of ZEB1 on *PBX3* may be indirectly mediated through this miRNA. Using TargetScan (Lewis et al.,

2005), we found that a 7-mer seed-matching sequence of *miR-200b/c* within in the *PBX3* 3'UTR was highly conserved across several species (Figure 43).

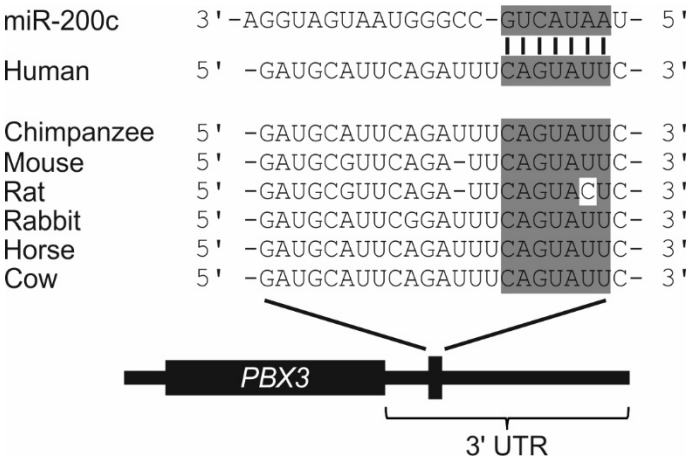

Figure 43: Putative miR-200c binding sites within the *PBX3* 3'UTR. Illustration of the miR-200c seeds and seed-matching sequences in the 3´-UTR of PBX3 (modified from www.targetscan.org).

Transfection of SW480 and Colo320 colon cancer cells with *miR-200c* repressed both ZEB1 and PBX3, with stronger effects on protein than on mRNA levels, as expected for direct miRNA effects (Figure 44).

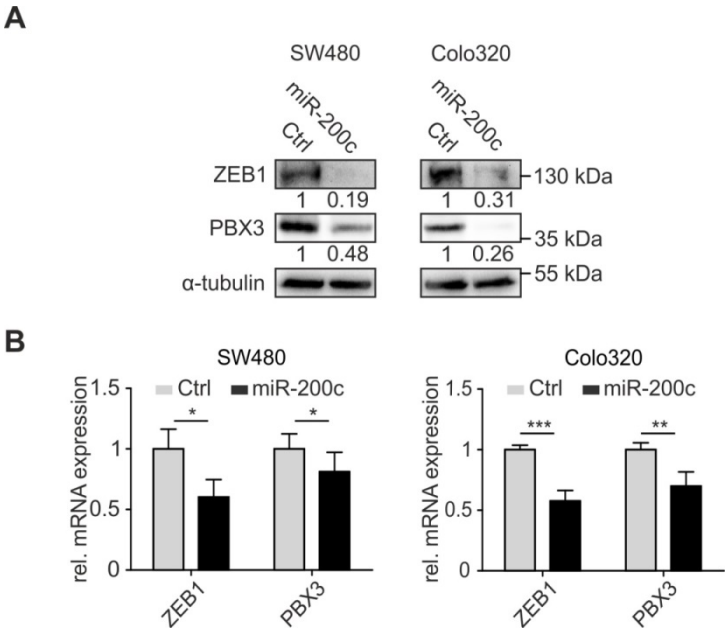

Figure 44: miR-200c represses PBX3 expression. (A) Western blot analysis of PBX3 protein levels in SW480 and Colo320 cells 48 hours after transfection with miR-200c or miR control. Numbers below immunoblots indicate fold change by densitometry. **(B)** PBX3 mRNA levels in the same experiment analogous to (A). Error bars indicate mean ± SD. *, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$ by t test, $n \geq 3$.

We then cloned the 3'UTR of *PBX3*, including the *miR-200* seed-matching sequence, 3' of a luciferase open reading frame (Figure 45), and found significant down-regulation of the reporter activity upon transfection with *miR-200c* (Figure 46), or upon siRNA mediated knockdown of ZEB1 (Figure 47).

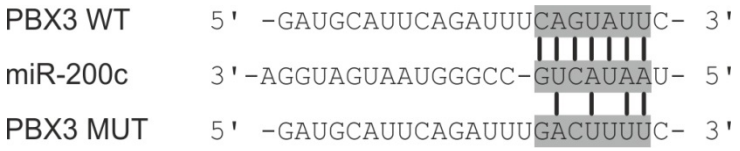

Figure 45: Mutation of the PBX3 3´-UTR. miR-200c seed and miR-200c seed-matching sequence shown with the remaining matches (black bars; WT: wild-type, MUT: mutated)

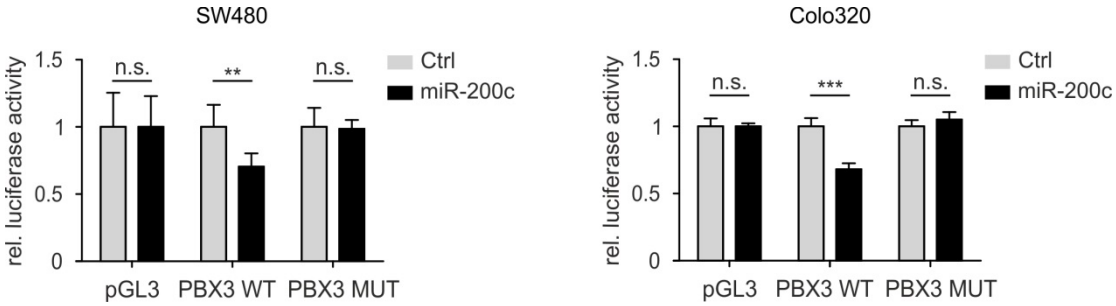

Figure 46: miR-200c represses the PBX3 reporter activity. Analysis of the luciferase reporter activity in SW480 and Colo320. Cells were transfected 72 hours with pre-miR-200c or control oligonucleotides and the empty pGL3 vector or pGL3 with the indicated 3´-UTR-reporter constructs. Error bars indicate mean ± SD. **, P < 0.01; ***, P < 0.001; n.s., not significant by t test, n ≥ 3.

Both effects were abolished when using a luciferase reporter containing the 3'UTR of *PBX3* with a mutated seed-matching sequence (Figures 46 and 47). These findings demonstrated that PBX3 is targeted by *miR-200c*, and suggested that ZEB1 mediated induction of PBX3 occurs indirectly through de-repression of *miR-200c*.

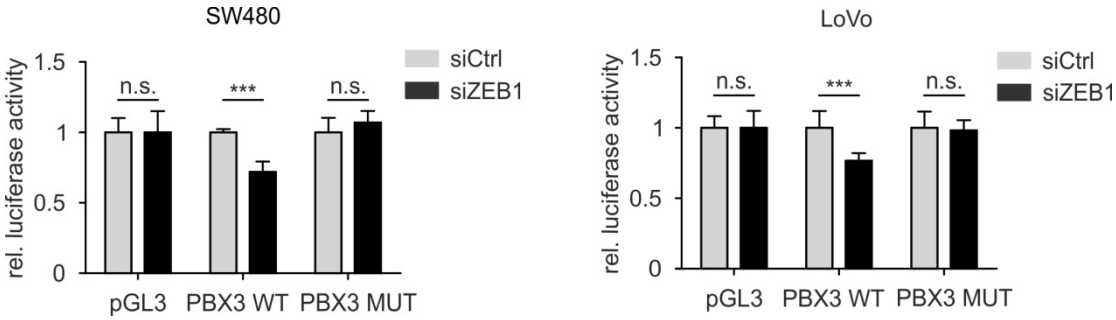

Figure 47: ZEB1 represses PBX3 reporter activity. Analysis of the luciferase reporter activity in SW480 and Colo320. Cells were transfected 72 hours with siRNA against ZEB1 or control

oligonucleotides and the empty pGL3 vector or pGL3 with the indicated 3´-UTR-reporter constructs. Error bars indicate mean ± SD., ***, P < 0.001; n.s., not significant by t test, n ≥ 3.

To further determine if EMT in colon cancer also depended on PBX3 expression, we induced EMT by SNAIL in DLD-1 and LS174T cells and concomitantly depleted PBX3 by siRNA. While SNAIL expression caused downregulation of E-Cadherin, indicating loss of epithelial features, depletion of PBX3 partially reversed this effect (Figure 48).

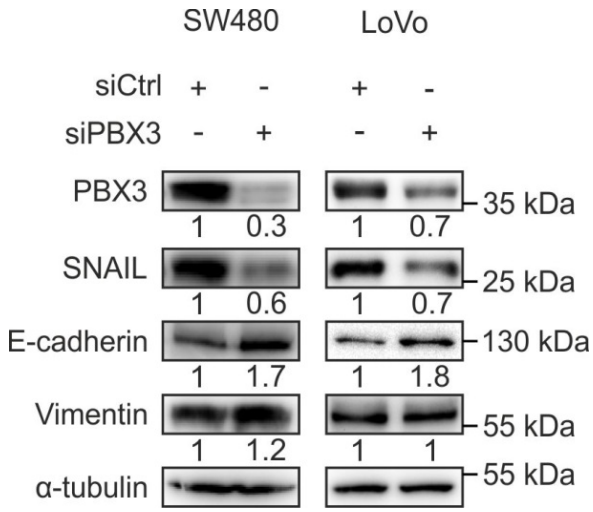

Figure 48: PBX3 affects EMT signaling. Immunoblotting of indicated proteins after transfection of SW480 and LOVO colon cancer cells with siRNA against PBX3 for 72 hours. Numbers below immunoblots indicate fold change by densitometry.

Moreover, depleting PBX3 in SW480 and LoVo cell lines caused upregulation of E-cadherin expression (Figure 49). These data implied that PBX3 expression is required for a full EMT phenotype in colon cancer cells.

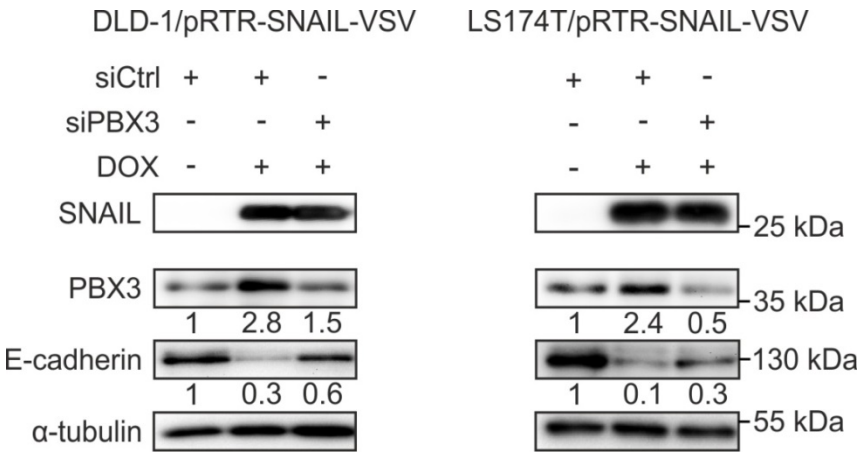

Figure 49: Knockdown of PBX3 inhibited the SNAIL dependent reduction of E-cadherin. Western blot analysis in DLD-1 and LS174T cells with the pRTR-SNAIL vector for the respective proteins. Cells were transfected for 48 hours with the indicated siRNAs and further stimulated with either DOX or left untreated for further 36 hours. Numbers below immunoblots indicate fold change by densitometry.

5.2.5. High PBX3 expression is a strong indicator of colon cancer progression

Due to its dependence on WNT signaling and EMT, both drivers of colon cancer progression, we examined the clinical significance of PBX3 expression in a collection of 244 colorectal cancer cases, all of which were stage II with clinical follow-up records (Table 3).

Table 3: Clinical data and PBX3 expression in UICC stage II colorectal cancer.

Characteristics	Total	PBX3 expression		*P*
		negative	positive	
All patients	244 (100)	54 (22.1)	190 (77.9)	
Age (y, median 69.3)				
≤ 69	122 (100.0)	19 (15.6)	103 (84.4)	0.014
≥ 70	122 (100.0)	35 (28.7)	87 (71.3)	
Gender				
Male	131 (100.0)	27 (20.6)	104 (79.4)	0.54
Female	113 (100.0)	27 (23.9)	86 (76.1)	
T-stage (UICC)				
T3	201 (100.0)	41 (20.4)	160 (79.6)	0.16
T4	43 (100.0)	13 (30.2)	30 (69.8)	
Tumor grade (WHO)				
low	146 (100.0)	29 (19.9)	117 (80.1)	0.3
high	98 (100.0)	25 (25.5)	73 (74.5)	

Row percent values are given in parentheses

While 72 tumors (30%) showed absence of PBX3 expression (score 0), 160 tumors (66%) had moderate levels (scores 1-2), and only 12 tumors (5%) high levels of PBX3 (score 3, Figure 50).

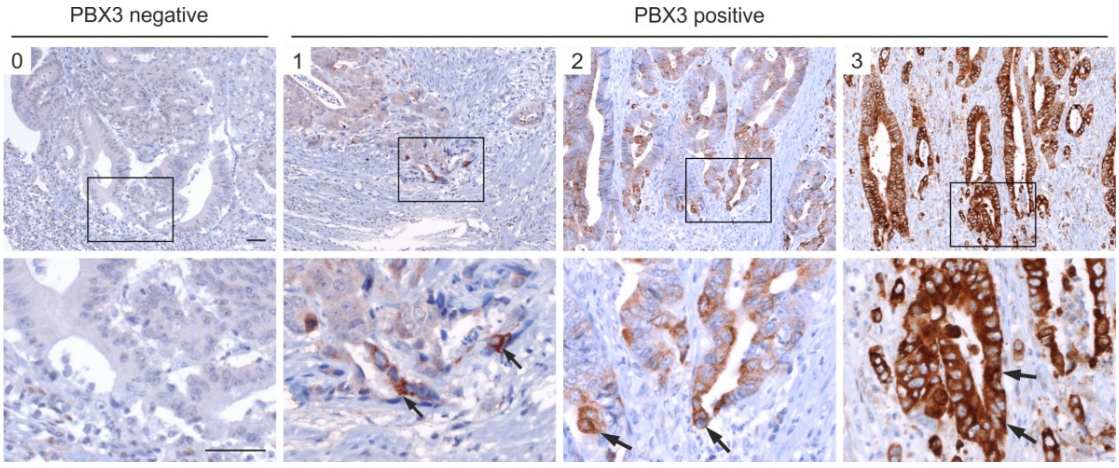

Figure 50: PBX3 expression in colorectal cancer indicates poor prognosis. Assessment of PBX3 immunostaining in primary human colorectal cancers. Tumors were selected semiquantitative expression scores from 0 (no staining) to 3 (strong staining) and accordingly categorized as PBX3 negative (score 0) and positive (scores 1-3). Arrows indicate stained tumor cells. Scale bar, 100 µm.

Kaplan-Meier statistics indicated significantly worse cancer specific patient survival when tumors expressed PBX3 (scores 1-3), when compared to lack of PBX3 expression (score 0). Importantly, when choosing tumor progression as endpoint, we found an even stronger correlation of PBX3 expression and poor prognosis (Figure 51). Both findings were independent from other clinical variables in proportional hazards regression analyses (Tables 4 and 5).

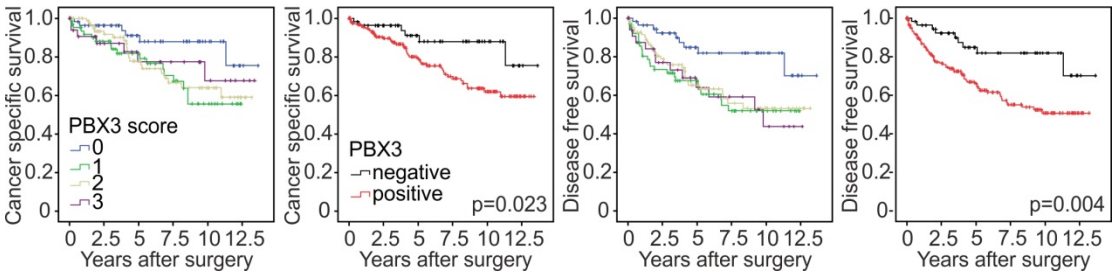

Figure 51: Loss of PBX3 expression reveals poor prognosis in CRC. Kaplan-Meier plots for different PBX3 expression scores and categories (lower panel) in a collection of n=244 stage II colorectal cancers.

Table 4: Multivariate analysis of cancer specific survival in UICC stage II colorectal cancer.

Variables	Cancer specific survival		
	HR	**(95% confidence interval)**	***P***
Age (≥ vs < median)	1.8	(1.02-3.09)	0.042
Gender (F vs M)	0.8	(0.45-1.35)	0.367
T-stage	2.5	(1.38-4.58)	0.003
Tumor grade	2.2	(1.15-4.07)	0.017
PBX3 positive vs negative	3.2	(1.35-7.61)	0.008

Table 5: Multivariate analysis of disease free survival in UICC stage II colorectal cancer.

Variables	Cancer specific survival		
	HR	**(95% confidence interval)**	***P***
Age (≥ vs < median)	1.1	(0.67-1.67)	0.816
Gender (F vs M)	0.8	(0.51-1.28)	0.366
T-stage	2.6	(1.58-4.29)	0.00017
Tumor grade	2.0	(1.20-3.40)	0.008
PBX3 positive vs negative	3.0	(1.50-6.14)	0.002

Since tumor outcome of colon cancer mainly depends on distant metastasis, we further investigated PBX3 expression in a second, independent matched case-control collection of 45 pairs of colon cancers with and without synchronous liver metastasis. In this collection, positive PBX3 expression (scores 1-3) were

significantly associated with liver metastasis (p=0.01, OR=3.0), further strengthening the link of PBX3 and poor prognosis (Table 6).

Table 6: Clinical data and PBX3 expression in a case-control collection of colon cancers with and without distant metastasis.

Characteristics	Total	PBX3 expression		*P*
		negative	positive	
All patients	90 (100)	42 (46.7)	48 (53.3)	
Age (y, Median 68)				
≤ 68	47 (52.2)	20 (42.6)	27 (57.4)	0.41
≥ 69	43 (47.8)	22 (51.2)	21 (48.8)	
Gender				
Male	44 (48.9)	18 (40.9)	26 (59.1)	0.28
Female	46 (51.1)	24 (52.2)	22 (47.8)	
T-stage (UICC)				
T2	8 (8.9)	2 (25)	6 (75	0.43
T3	66 (73.3)	32 (48.5)	34 (51.5)	
T4	16 (17.8)	8 (50)	8 (50)	
Nodal status				
N0	38 (42.2)	19 (50)	19 (50)	0.58
N+	52 (57.8)	23 (44.2)	29 (55.8)	
Metastasis (Liver)				
M0	45 (50)	27 (60)	18 (40)	0.011
M1	45 (50)	15 (33.3)	30 (66.7)	
Tumor grade (WHO)				
Low	30 (33.3)	8 (26.7)	22 (73.3)	0.007
High	60 (66.7)	34 (56.7)	26 (43.3)	

Row percent values are given in parentheses

Finally, for independent validation of these results, we analyzed clinical correlations of *PBX3* mRNA levels with the assembled gene expression data set of 1,032 colon cancer cases, 927 of which had follow-up data on tumor progression. Using ROC curve analyses, we identified an ideal cutoff at the normalized expression intensity of 277 (natural scale) of *PBX3* mRNA. Dichotomal classification of cases by this score revealed a highly significant positive correlation of high *PBX3* expression and tumor progression in this data set (Figure 52). Collectively, PBX3 is a strong prognostic marker for tumor progression and poor survival in patients with colorectal cancer.

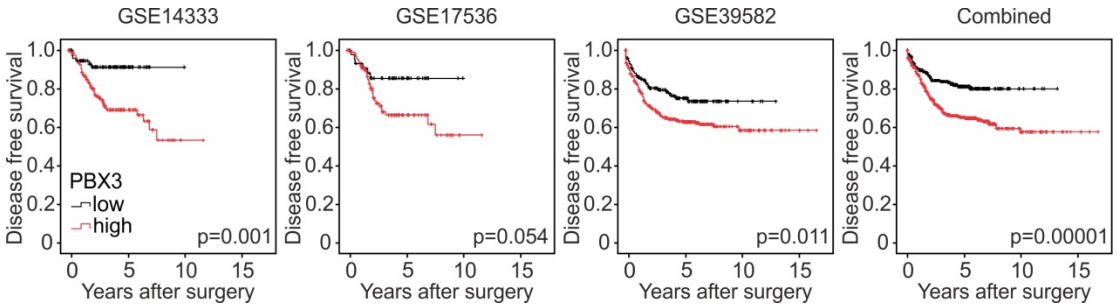

Figure 52: PBX3 expression in colorectal cancer indicates poor prognosis. PBX3 mRNA expression and survival association in individual and combined data sets of a total of n=927 colon cancers. GEO accession numbers of individual data sets are indicated. Kaplan-Meier plots for cases with low and high PBX3 expression. P values are log-rank test results.

6. Discussion

6.1. Multicolor lineage tracing reveals clonal architecture and dynamics in colon cancer

In this study, we implemented a quantitative lineage tracing strategy to gain unbiased insights into the clonal expansion dynamics of individual tumor cells within growing colon cancer *in vivo*. We used cell line derived colon cancer xenografts as model tumors to reproduce the architecture, cellular composition, and differentiation of primary human colon cancers (Cernat et al., 2014). Our data illustrate that colon cancer cells at the leading tumor edge compete for clonal outgrowth which is directed towards the tumor necrotic core. Although most colon cancers show central necrosis, this probably is not due to a commitment of colon cancer cells to die but rather caused by insufficient nutrient supply in the center of rapidly growing tumors.

In tumors with organized differentiation gradients this clonal expansion may coincide with tumor cell differentiation. These findings are in agreement with recent data demonstrating clonal outgrowth from tumor cells with high MAPK pathway activity or high expression of the WNT target gene LGR5 at the leading tumor edge (Blaj et al., 2017; Shimokawa et al., 2017). In this case, linear expansion of tumor cell subclones may be well compatible with lineage outgrowth from phenotypically defined colon cancer stem cells (Clevers, 2011). Despite a distorted architecture, clonal outgrowth and differentiation in colon cancer therefore can be reminiscent of normal colonic mucosa, where stem cells at the crypt base compete for clonal repopulation of individual crypts (Chan et al., 2009; Snippert et al., 2010). Linear expansion of colon cancer subclones may thus be well compatible with lineages of a cancer stem cell model (Reya et al., 2001).

However, colon cancer xenografts lacking differentiation gradients, and thus more disorganized growth, unexpectedly showed the same pattern of clonal expansion from tumor edge to center. Therefore, clonal fitness and positive clonal selection rather appear to depend on positioning of tumor cells at the leading tumor edge than on tumor cell differentiation. Indeed, when

considering widespread expression of putative cancer stem cell antigens in colon cancers (Horst et al., 2012), it may be difficult to imagine how a putative cancer stem cell that is trapped centrally within the tumor mass should efficiently compete for space and resources required for clonal expansion (Greaves and Maley, 2012). Based on our data, and supported by the results of our spatial computer model, which implemented positon as the only factor determining tumor cell behavior, we therefore propose that competition of colon cancer cells for clonal expansion is mainly restricted to the leading tumor edge. The phenotype of tumor cells within expanding clones may still be variable, depend on the individual genetic background of the tumor, and may secondarily be influenced by a position-related tumor microenvironment (Brabletz et al., 2001).

Previous attempts to follow individually labelled tumor cells over time, independently of their phenotype, either used murine models, or lentiviral color- or bar-coding methods for random genetic labeling of tumor cells *in vitro* before xenotransplantation into mice (Cornils et al. 2014; Dieter et al. 2011; Weber et al. 2011; Zomer et al. 2013). While the approach in murine models is not applicable to human malignancies, the *in vitro* labeling approach has the caveat that clonal cell tracing cannot be induced after secondary tumor architectures have formed, and thus precludes access to clonal fate data of individual tumor cells. By combining the advantages of inducible clonal cell tracing and lentiviral delivery, we overcome these restrictions, and for the first time demonstrate a constant drift towards oligoclonality within colon cancer that appears to be based on clonal competition and axial outgrowth.

Xenograft growth of human colon cancer cell lines seems to be the appropriate model as the architecture and tumor cell heterogeneity of primary human colon cancers are properly reflected. By contrast, for studies addressing metastasis formation orthotopic tumor implantations would be the adequate model. However, in our study we examined clonal outgrowth of xenografts over relatively long periods of time (31 days) so that tumors reach diameters of up to 1.5 cm, which would be in conflict with orthotopic tumor implantation.

However, although we simulate these dynamics in our computer model by neutral stochastic competition of tumor cells at the leading tumor edge, the biological basis for clonal competition yet remains to be determined. Also, due to a limited number of different fluorescent colors, some aspects including the significance of neighboring clones with identical colors or clone fragmentation during outgrowth may be missed by our labelling strategy and require further study.

Importantly, in our model this clonal competition does not depend on mutational evolution. Although additionally acquired mutations in individual tumor cell subclones may provide fitness advantage, genetic changes that substantially alter the clonal composition of a final tumor are assumed to be rare events in rapidly expanding cell populations (Korolev et al., 2010; McFarland et al., 2013). An inferred “Big Bang” model of colon cancer evolution therefore suggested that clonal dynamics in established tumors are mainly devoid of substantial mutational evolution (Sottoriva et al., 2015). In line with this idea, we found no differences in driver mutation profiles of individual tumor cell subclones. Therefore, we suggest that clonal competition in colon cancer is mainly determined by tumor cell positon, and may continuously occur throughout the lifespan of a tumor.

Nevertheless, it remains to be determined to what extent other heritable traits may have an impact on clonal architecture and growth dynamics, since others reported epigenetic differences among subclones of colonic adenomas and colon cancers (Humphries et al., 2013; Siegmund et al., 2009). However, in contrast to unperturbed tumor growth, mutational evolution certainly plays an important role in acquired resistance to targeted therapy (Diaz Jr et al., 2012). Previous data suggested that treatment protocols stabilizing tumor growth rather than attempting to eradicate the tumor may prolong cancer survival (Gatenby et al., 2009). Our data suggesting continuous clonal competition may explain such findings. If treatment-resistant tumor subclones have to compete for space and resources with treatment-sensitive subclones that prevail under gentle targeted therapy, loss of resistant clones into tumor necrosis may occur by chance. In contrast, harsher targeted therapy may eliminate sensitive tumor cell clones and strongly favor a quick outgrowth of

resistant clones with earlier treatment failure. This hypothesis however, will require further experimental proof, and may then inform the design of future targeted therapeutic approaches for colon cancer patients.

6.2. PBX3 is part of an EMT regulatory network in colorectal cancer and indicates poor outcome

The ability of epithelial cancer cells to loose cellular junctions and polarity with subsequent infiltration of tumor surrounding stromal tissue is a main aspect of EMT and hallmark of cancer progression (Hanahan and Weinberg, 2011). Here we identify strong overexpression of PBX3 in tumor cells with high WNT activity undergoing EMT at the leading tumor edge of colorectal cancers. We demonstrate that PBX3 expression is induced in this tumor cell subset by WNT and the EMT regulating transcription factors SNAIL and ZEB1, while this induction – at least partially – occurs indirectly through a decreased repression of *PBX3* mRNA by *miR-200*. These findings are in agreement with recent data that demonstrated targetability of *PBX3* by different microRNAs (Han et al., 2015), and therefore place its expression in colon cancer downstream of a WNT and EMT regulatory network (Thiery et al., 2009). Furthermore, we demonstrate that PBX3 expression is required for a full EMT phenotype in colon cancer cells, since its depletion partially blocked EMT induction by ZEB1 and SNAIL, and increased the expression of E-cadherin, indicating a shift towards more epithelial differentiation. In line with this finding PBX3 has recently been shown to induce EMT in gastric cancer cells (Han et al., 2014; Li et al., 2017), indicating that it may generally be involved in EMT regulation in gastrointestinal cancers. Because PBX3 also has been shown to increase migration and invasion of colon cancer cells (Han et al., 2014) both of which are phenotypic characteristics of tumor cells undergoing EMT, this further supports the notion that PBX3 directly contributes to the infiltrative phenotype of colon cancer cells at the leading tumor edge. However, the exact mechanism by which PBX3 influences EMT in colon cancer still remains to be determined, keeping in mind that PBX3 may function as cofactor for homeobox proteins (Li et al., 2013).

In primary colon cancer tissues, we found that PBX3 can easily be visualized *in situ* by immunostaining. Importantly, labelling in these tumors was restricted to cancer cells while tumor surrounding stromal cells were PBX3 negative. Given that markers which robustly indicate EMT in colon cancer are scarce,

and detection of ZEB1, SNAIL and Vimentin can be difficult and confounded by labelling of stromal cells (Francí et al., 2009; Spaderna et al., 2006; Toiyama et al., 2013), we propose that PBX3 may be a useful marker to highlight and further study colon cancer cells undergoing EMT *in situ*. Furthermore, we demonstrate that PBX3 mRNA levels strongly correlated with EMT in a large gene expression data set derived from 1,032 colon cancer samples. Considering the restriction of PBX3 expression to cancer cells, we therefore propose that on the gene expression level PBX3 may indicate the overall degree of EMT in colon cancer specimens with little confounding by the amount of stromal tissue within each sample. Of note however, PBX3 expression was not completely restricted to infiltrative tumor cells at the leading tumor edge but also extended to glandular differentiated colon cancer cells, especially in cases with high levels of PBX3 expression. Because similar observations also were made for ZEB1 and SNAIL (Francí et al., 2009; Spaderna et al., 2006), it remains to be determined to what extent infiltrative tumor cell morphology and EMT related factors indicate identical or only partially overlapping colon cancer cell subpopulations.

Our findings in case collections with clinical follow-up data demonstrate that PBX3 expression is strongly linked to poor outcome in patients with colorectal cancer. High PBX3 expression was significantly associated with poor cancer specific survival and strongly correlated with an increased risk for cancer progression in a collection of 244 stage II colorectal cancers, while this was independent of other core clinical variables. Stage II colorectal cancer is characterized by local disease with full-thickness involvement of the bowel wall but absence of lymphatic or distant metastasis (Edge and Compton, 2010). Accordingly, most of these patients can be cured by surgical resection alone. However, disease progression after surgery still is observed in 25-30% of these cases and patients may eventually die from their disease (Dotan and Cohen, 2011). We therefore suggest that assessing PBX3 expression may identify potentially aggressive cases of stage II colorectal cancer that may benefit from adjuvant chemotherapy despite low clinical stage (Gunderson et al., 2010). Integrating PBX3 expression analysis into routine pathology workup of colorectal cancer specimens thus may guide the decision for

therapeutic management in addition to staging. However, PBX3 expression also strongly correlated with metastasis in our case control collection and thus also indicated disease progression in late stage disease. Moreover, we found that on the mRNA level PBX3 expression was highly significantly associated with poor outcome in a combined dataset with clinical information on 923 colon cancers including all stages. This not only further confirmed our findings but also validated results from a previous study that suggested an association of PBX3 mRNA expression and poor patient survival in a smaller case collection (Han et al., 2014). Taken together, we here establish PBX3 as a robust marker for outcome stratification in patients with colorectal cancer. Due to its association with EMT, which regulates invasion and metastasis as a basis for cancer progression, we suggest that the independent prognostic power of PBX3 may be due to gauging EMT which is not sufficiently reflected by other clinical and pathological variables.

7. Summary

Colon cancers are composed of phenotypically heterogeneous tumor cell subpopulations with variable expression of putative stem cell and differentiation antigens. While in normal colonic mucosa clonal repopulation occurs among differentiation gradients from crypt base towards crypt apex, the clonal architecture of colon cancer and the relevance of tumor cell subpopulations for clonal outgrowth are poorly understood.

In the first study of this thesis, we used a multicolor lineage tracing approach in colon cancer xenografts that reflect primary colon cancer architecture. With this method, we could demonstrate that clonal outgrowth is mainly driven by tumor cells located at the leading tumor edge with clonal axis formation towards the tumor center. Our findings suggest that in colon cancer tumor cell position may be more important for clonal outgrowth than tumor cell phenotype.

In a second study we analyzed colon cancer cells with high WNT signaling activity. This characteristic overactivation in colorectal cancers is caused by pathway activating mutations and drives tumor progression and metastasis. Here, we could identify pre-leukemia transcription factor 3 (PBX3) as a gene regulated in part by WNT signaling in colon cancers and assess its prognostic value. In a colon cancer case collection, PBX3 expression correlated with nuclear β-catenin and high PBX3 levels were associated with decreased patient survival and an increased risk for tumor relapse and metastasis. Additionally, an independent case control study confirmed the association of high PBX3 expression and colon cancer metastasis to the liver. Further studies provide evidence that PBX3 is also regulated by EMT. Knockdown and overexpression studies modifying the EMT transcriptions factors SNAIL and ZEB1 as well as PBX3 demonstrated that PBX3 is part of an EMT-regulatory network in colorectal cancer.

Taken together the results of these two studies may have implications for the cancer stem cell hypothesis, reveal new aspects of tumor cell heterogeneity in colorectal cancer, and could furthermore give insights for the design of new therapeutic strategies.

8. Zusammenfassung

Dickdarmkarzinome bestehen aus phänotypisch heterogenen Subpopulationen von Tumorzellen, mit unterschiedlicher Expression putativer Stammzell- und Differenzierungsantigene. Während in normaler Dickdarmschleimhaut die klonale Besiedelung entlang Differenzierungsgradienten von der Basis der Krypte zur Spitze hin erfolgt, ist die klonale Architektur bei Dickdarmkarzinomen und die Bedeutung von Tumorzellsubpopulationen in Bezug auf den klonalen Auswuchs wenig bekannt.

In der ersten Studie untersuchten wir Dickdarmkarzinom-Xenotransplantate, die die primäre Architektur des Kolonkarzinoms widerspiegeln, mithilfe eines multifarben „lineage tracing“ Ansatzes. Anhand dieser Methode konnten wir zeigen, dass Tumorzellen am Tumorrand hauptsächlich für das klonale Wachstum im Tumor verantwortlich sind und sich dieses entlang einer Achse zum Tumorzentrum hin erstreckt. Diese Ergebnisse zeigen, dass im Dickdarmkarzinom für das klonale Wachstum die Position der Tumorzellen offenbar wichtiger ist als deren Phänotyp.

In einer zweiten Studie untersuchten wir Tumorzellen mit hoher WNT Signalwegaktivität. Diese für Kolonkarzinome charakteristische Überaktivierung entsteht durch den Signalweg aktivierende Mutationen und trägt zur Tumorprogression und Metastasierung bei. Hierbei entdeckten wir, dass Prä-Leukämie Transkriptionsfaktor 3 (PBX3) im Kolonkarzinom zum Teil durch den WNT Signalweg reguliert wird. In einer Fallsammlung von Dickdarmkarzinomen überlappte die PBX3 Expression mit nukleärem β-catenin und hohe Expressionslevel von PBX3 konnten mit einem geringeren Patientenüberleben sowie mit erhöhtem Risiko eines Tumorrezidivs und Metastasenbildung assoziiert werden. Des Weiteren konnte eine unabhängige Fallstudie den Zusammenhang zwischen hoher PBX3 Expression und Metastasen des Dickdarmkarzinoms bestätigen. Weitere Experimente lieferten den Beweis, dass PBX3 zudem durch EMT reguliert wird. RNA Interferenz und Überexpressionsstudien bei denen die EMT

Transkriptionsfaktoren SNAIL und ZEB1 sowie PBX3 modifiziert wurden, zeigten, dass PBX3 Teil eines EMT-regulatorischen Netzwerks beim Dickdarmkarzinom ist.

Zusammenfassend haben die Ergebnisse dieser Studien Auswirkungen auf die Stammzellhypothese und zeigen neue Aspekte der Tumorzellheterogenität beim Dickdarmkrebs auf. Darüberhinaus können sie für die zukünftige Entwicklung neuer therapeutischer Strategien hilfreiche Einblicke geben.

9. Abbreviations

AML	acute myeloid leukemia
APC	adenomatous polyposis coli
APS	ammonium peroxodisulfate
BFP	blue fluorescence protein
BlastR	blasticidin resistance
Bmi1	B lymphoma Mo-MLV insertion region 1 homolog
bp	basepairs
BrdU	Bromodeoxyuridine
BSA	bovine serum albumin
CBC	crypt based columnar cell
CDK	cyclin-dependent kinase
cDNA	complementary DNA
CE	clonal evolution
CIN	chromosomal instability
CK1	casein kinase 1
CRC	colorectal cancer
CSC	cancer stem cell
CFP	cyan fluorescent protein
DAPI	2-(4-amidinophenyl)-6-indolecarbamidine dihydrochloride
DMEM	Dulbecco`s modified Eagles medium
DMSO	dimethyl-sulfoxide
DNA	deoxyribonucleic acid
dnTCF4	dominant negative TCF4
dNTP	deoxynucleotides triphosphate
DOX	doxycycline
Dvl	dishevelled
E.coli	*Escherichia coli*
EGFR	epidermal growth factor receptor
EMT	epithelial-mesenchymal transition
FBS	fetal bovine serum
FoxC2	Forkhead box protein C2
GSEA	Gene Set Enrichment Analyses
GSK3	glycogen synthase kinase 3
HCC	hepatocellular carcinoma
Hopx	HOP homeobox
HRP	horseradish peroxidase
IF	immunofluorescence
IHC	immunohistochemistry
int-1	integrated 1
KLF8	krueppel-like factor 8
LB	lysogeny both

LEF1	lymphoid enhancer-binding factor 1
let-7	lethal-7
Lgr5	leucine rich repeat containing G protein coupled receptor 5
LRP 5/6	lipoprotein receptor related protein 5/6
LTR	long terminal repeat
MAPK	mitogen-activated protein kinase
MET	mesenchymal-epithelial transition
miR	microRNA
MMR	mismatch repair
mRNA	messenger RNA
MSI	microsatellite instability
mut	mutated
NIH	National Institutes of Health
OFP	orange fluorescent protein
ORF	open reading frame
PBX	pre-B-cell leukemia homeobox
PBS	phosphate buffered saline
(q)PCR	(quantitative) polymerase chain reaction
PI3K	phosphoinositide-3-kinase
pri-miR	primary microRNA transcript
PuroR	puromycin resistance
RFP	red fluorescent protein
RNA	ribonucleic acid
RT	room temperature
RUNX3	runt-related transcription factor 3
SD	standard deviation
SDS	sodium dodecyl sulfate
TAM	tamoxifen
TCF	T cell transcription factor
Temed	tetramethylethylenediamine
Tert	telomerase reverse transcriptase
Tet	tetracycline
TF	transcription factor
TGFB1	transforming growth factor –β
TIC	tumor-initiating cell
TMA	tissue microarray
TRE	tetracycline response element
TRS	target Retrieval Solution
TSP1	thrombospondin 1
TSS	transcription start site
RFP	red fluorescent protein
rtTa	reverse tetracycline-controlled transactivator
UTR	untranslated region
VSV	vesicular stomatitis virus (tag)

WB	western blot
WT	wild-type
XFP	fluorescent protein
YFP	yellow fluorescent protein
ZEB	zinc finger E-box-binding homeobox protein

10. References

Aberle, H., Bauer, A., Stappert, J., Kispert, A., and Kemler, R. (1997). Beta-Catenin Is a Target for the Ubiquitin-Proteasome Pathway. EMBO J. *16*, 3797–3804.

Amit, S., Hatzubai, A., Birman, Y., Andersen, J.S., Ben-Shushan, E., Mann, M., Ben-Neriah, Y., and Alkalay, I. (2002). Axin-mediated CKI phosphorylation of β-catenin at Ser 45: A molecular switch for the Wnt pathway. Genes Dev. *16*, 1066–1076.

Anastassiou, D., Rumjantseva, V., Cheng, W., Huang, J., Canoll, P.D., Yamashiro, D.J., and Kandel, J.J. (2011). Human cancer cells express Slug-based epithelial-mesenchymal transition gene expression signature obtained in vivo. BMC Cancer *11*, 529.

Arnoux, V., Nassour, M., L'Helgoualc'h, A., Hipskind, R.A., and Savagner, P. (2008). Erk5 controls Slug expression and keratinocyte activation during wound healing. Mol. Biol. Cell *19*, 4738–4749.

Bao, B., Ahmad, A., Azmi, A.S., Ali, S., and Sarkar, F.H. (2013). Overview of cancer stem cells (CSCs) and mechanisms of their regulation: implications for cancer therapy. Curr. Protoc. Pharmacol. *Chapter 14*, Unit 14.25.

Barker, N. (2014). Adult intestinal stem cells: critical drivers of epithelial homeostasis and regeneration. Nat. Rev. Mol. Cell Biol. *15*, 19–33.

Barker, N., van Es, J.H., Kuipers, J., Kujala, P., van den Born, M., Cozijnsen, M., Haegebarth, A., Korving, J., Begthel, H., Peters, P.J., et al. (2007). Identification of stem cells in small intestine and colon by marker gene Lgr5. Nature *449*, 1003–1007.

Barker, N., Van Oudenaarden, A., and Clevers, H. (2012). Identifying the stem cell of the intestinal crypt: Strategies and pitfalls. Cell Stem Cell *11*, 452–460.

Batlle, E., Sancho, E., Francí, C., Domínguez, D., Monfar, M., Baulida, J., and García De Herreros, a (2000). The transcription factor snail is a repressor of E-cadherin gene expression in epithelial tumour cells. Nat. Cell Biol. *2*, 84–89.

Bjerknes, M., and Cheng, H. (1981). The stem cell zone of the small intestinal epithelium. I. Evidence from paneth cells in the adult mouse. Am. J. Anat. *160*, 51–63.

Blaj, C., Schmidt, E.M., Lamprecht, S., Hermeking, H., Jung, A., Kirchner, T., and Horst, D. (2017). Oncogenic effects of high MAPK activity in colorectal cancer mark progenitor cells and persist irrespective of RAS mutations. Cancer Res. *77*, 1763–1774.

Blanpain, C., Horsley, V., and Fuchs, E. (2007). Epithelial Stem Cells: Turning over New Leaves. Cell *128*, 445–458.

Bonnet, D., and Dick, J.E. (1997). Human acute myeloid leukemia is organized as a hierarchy that originates from a primitive hematopoietic cell. Nat Med *3*, 730–737.

Bosman, F.T., Carneiro, F., Hruban, R.H., and Theise, N.D. (2010). WHO Classification of Tumours of the Digestive System, Fourth Edition. In International Agency for Research on Cancer, p. 417.

Botteri, E., Iodice, S., Bagnardi, V., Raimondi, S., Lowenfels, A.B., and Maisonneuve, P. (2008). Smoking and colorectal cancer: a meta-analysis. JAMA *300*, 2765–2778.

Brabletz, T. (2012). To differentiate or not — routes towards metastasis. Nat. Rev. Cancer *12*, 425–436.

Brabletz, T., Jung, a, Reu, S., Porzner, M., Hlubek, F., Kunz-Schughart, L. a, Knuechel, R., and Kirchner, T. (2001). Variable beta-catenin expression in colorectal cancers indicates tumor progression driven by the tumor environment. Proc. Natl. Acad. Sci. U. S. A. *98*, 10356–10361.

Brabletz, T., Jung, A., Spaderna, S., Hlubek, F., and Kirchner, T. (2005). Opinion - Migrating cancer stem cells - an integrated concept of malignant tumour progression. Nat. Rev. Cancer *5*, 744–749.

Bristow, R.G., and Hill, R.P. (2008). Hypoxia and metabolism. Hypoxia, DNA repair and genetic instability. Nat. Rev. Cancer *8*, 180–192.

Buczacki, S.J. a, Zecchini, H.I., Nicholson, A.M., Russell, R., Vermeulen, L., Kemp, R., and Winton, D.J. (2013). Intestinal label-retaining cells are secretory precursors expressing Lgr5. Nature *495*, 65–69.

Cabrera, M.C., Hollingsworth, R.E., and Hurt, E.M. (2015). Cancer stem cell plasticity and tumor hierarchy. World J. Stem Cells *7*, 27–36.

Campbell, P.J., Yachida, S., Mudie, L.J., Stephens, P.J., Pleasance, E.D., Stebbings, L.A., Morsberger, L.A., Latimer, C., McLaren, S., Lin, M.-L., et al. (2010). The patterns and dynamics of genomic instability in metastatic pancreatic cancer. Nature *467*, 1109–1113.

Cano, a, Pérez-Moreno, M. a, Rodrigo, I., Locascio, a, Blanco, M.J., del Barrio, M.G., Portillo, F., and Nieto, M. a (2000). The transcription factor snail controls epithelial-mesenchymal transitions by repressing E-cadherin expression. Nat. Cell Biol. *2*, 76–83.

Cernat, L., Blaj, C., Jackstadt, R., Brandl, L., Engel, J., Hermeking, H., Jung, A., Kirchner, T., and Horst, D. (2014). Colorectal cancers mimic structural organization of normal colonic crypts. PLoS One *9*.

Chan, C.W.M., Wong, N. a, Liu, Y., Bicknell, D., Turley, H., Hollins, L., Miller, C.J., Wilding, J.L., and Bodmer, W.F. (2009). Gastrointestinal differentiation marker Cytokeratin 20 is regulated by homeobox gene CDX1. Proc. Natl. Acad. Sci. U. S. A. *106*, 1936–1941.

Cheng, H., and Leblond, C.P. (1974). Origin, differentiation and renewal of the four main epithelial cell types in the mouse small intestine I. Columnar cell. Am. J. Anat. *141*, 461–479.

Clevers, H. (2006). Wnt/beta-Catenin Signaling in Development and Disease. Cell *127*, 469–480.

Clevers, H. (2011). The cancer stem cell: premises, promises and challenges. Nat. Med. *17*, 313–319.

Clevers, H., and Nusse, R. (2012). Wnt/ β-catenin signaling and disease. Cell *149*, 1192–1205.

Cornils, K., Thielecke, L., Hüser, S., Forgber, M., Thomaschewski, M., Kleist, N., Hussein, K., Riecken, K., Volz, T., Gerdes, S., et al. (2014). Multiplexing clonality: Combining RGB marking and genetic barcoding. Nucleic Acids Res. *42*.

Crijns, A.P.G., de Graeff, P., Geerts, D., ten Hoor, K.A., Hollema, H., van der Sluis, T., Hofstra, R.M.W., de Bock, G.H., de Jong, S., van der Zee, A.G.J., et al. (2007). MEIS and PBX homeobox proteins in ovarian cancer. Eur. J. Cancer *43*, 2495–2505.

Diaz Jr, L. a., Williams, R.T., Wu, J., Kinde, I., Hecht, J.R., Berlin, J., Allen, B., Bozic, I., Reiter, J.G., Nowak, M. a., et al. (2012). The molecular evolution of acquired resistance to targeted EGFR blockade in colorectal cancers. Nature *486*, 4–7.

Dick, J.E. (2009). Looking ahead in cancer stem cell research. Nat. Biotechnol. *27*, 44–46.

Dieter, S.M., Ball, C.R., Hoffmann, C.M., Nowrouzi, A., Herbst, F., Zavidij, O., Abel, U., Arens, A., Weichert, W., Brand, K., et al. (2011). Distinct types of tumor-initiating cells form human colon cancer tumors and metastases. Cell Stem Cell *9*, 357–365.

Dotan, E., and Cohen, S.J. (2011). Challenges in the management of stage II colon cancer. Semin. Oncol. *38*, 511–520.

Driessens, G., Beck, B., Caauwe, A., Simons, B.D., and Blanpain, C. (2012). Defining the mode of tumour growth by clonal analysis. Nature *488*, 527–530.

Edge, S.B., and Compton, C.C. (2010). The American Joint Committee on Cancer: the 7th Edition of the AJCC Cancer Staging Manual and the Future of TNM. Ann. Surg. Oncol. *17*, 1471–1474.

Ellis-Connell, A.L., Iempridee, T., Xu, I., and Mertz, J.E. (2010). Cellular microRNAs 200b and 429 regulate the Epstein-Barr virus switch between latency and lytic replication. J. Virol. *84*, 10329–10343.

Fearon, E.R. (2011). Molecular genetics of colorectal cancer. Annu. Rev. Pathol. *6*, 479–507.

Fearon, E.R., and Vogelstein, B. (1990). A genetic model for colorectal tumorigenesis. Cell *61*, 759–767.

De Ferrari, G. V, and Moon, R.T. (2006). The ups and downs of Wnt signaling in prevalent neurological disorders. Oncogene *25*, 7545–7553.

Francí, C., Gallén, M., Alameda, F., Baró, T., Iglesias, M., Virtanen, I., and de Herreros, A.G. (2009). Snail1 protein in the stroma as a new putative prognosis marker for colon tumours. PLoS One *4*.

Gammons, M. V., Renko, M., Johnson, C.M., Rutherford, T.J., and Bienz, M. (2016). Wnt Signalosome Assembly by DEP Domain Swapping of Dishevelled. Mol. Cell *64*, 92–104.

Gatenby, R.A., Silva, A.S., Gillies, R.J., and Frieden, B.R. (2009). Adaptive therapy. Cancer Res. *69*, 4894–4903.

Gerdes, M.J., Sood, A., Sevinsky, C., Pris, A.D., Zavodszky, M.I., and Ginty, F. (2014). Emerging understanding of multiscale tumor heterogeneity. Front. Oncol. *4*, 366.

Giovannucci, E. (2002). Modifiable risk factors for colon cancer. Gastroenterol. Clin. North Am. *31*, 925–943.

Gonzalez, D.M., and Medici, D. (2014). Signaling mechanisms of the epithelial-mesenchymal transition. Sci. Signal. *7*, re8-re8.

Grady, W.M., Rajput, A., Myeroff, L., Liu, D.F., Kwon, K., Willis, J., and Markowitz, S. (1998). Mutation of the type II transforming growth factor-β receptor is coincident with the transformation of human colon adenomas to malignant carcinomas. Cancer Res. *58*, 3101–3104.

Grady, W.M., and Markowitz, S.D. (2008). TGF-β Signaling Pathway and Tumor Suppression. In Cold Spring Harbor Monograph Archive; Volume 50 (2008): The TGF-β Family, pp. 889–937.

Greaves, M., and Maley, C.C. (2012). Clonal evolution in cancer. Nature *481*, 306–313.

Grünewald, T.G.P., Bernard, V., Gilardi-hebenstreit, P., Surdez, D., Aynaud, M., Mirabeau, O., Cidre-, F., Tirode, F., Zaidi, S., Perot, G., et al. (2016). HHS Public Access. *47*, 1073–1078.

Gunderson, L.L., Jessup, J.M., Sargent, D.J., Greene, F.L., and Stewart, A.K. (2010). Revised TN categorization for colon cancer based on national survival outcomes data. J. Clin. Oncol. *28*, 264–271.

Hahn, S., and Hermeking, H. (2014). ZNF281/ZBP-99: A new player in epithelial-mesenchymal transition, stemness, and cancer. J. Mol. Med. *92*, 571–581.

Hahn, S., Jackstadt, R., Siemens, H., Hünten, S., and Hermeking, H. (2013). SNAIL and miR-34a feed-forward regulation of ZNF281/ZBP99 promotes epithelial-mesenchymal transition. EMBO J. *32*, 3079–3095.

Han, H., Gu, J., Zuo, H., Chen, Z., Zhao, W., Li, M., Ji, D., and Lu, Y. (2012). Let-7c functions as a metastasis suppressor by targeting MMP11 and PBX3 in colorectal cancer. 544–555.

Han, H., Du, Y., Zhao, W., Li, S., Chen, D., Zhang, J., Liu, J., Suo, Z., Bian, X., Xing, B., et al. (2015). PBX3 is targeted by multiple miRNAs and is essential for liver tumour-initiating cells. Nat. Commun. *6*, 8271.

Han, H.-B., Gu, J., Ji, D.-B., Li, Z.-W., Zhang, Y., Zhao, W., Wang, L.-M., and Zhang, Z.-Q. (2014). PBX3 promotes migration and invasion of colorectal cancer cells via activation of MAPK/ERK signaling pathway. World J. Gastroenterol. *20*, 18260–18270.

Hanahan, D., and Weinberg, R.A. (2000). The hallmarks of cancer. Cell *100*, 57–70.

Hanahan, D., and Weinberg, R.A. (2011). Hallmarks of cancer: The next generation. Cell *144*, 646–674.

Hay, E.D. (1995). An Overview of Epithelio-Mesenchymal Transformation. Cells Tissues Organs *154*, 8–20.

He, T.C., Chan, T.A., Vogelstein, B., and Kinzler, K.W. (1999). PPARdelta is an APC-regulated target of nonsteroidal anti-inflammatory drugs. Cell *99*, 335–345.

Horst, D., Chen, J., Morikawa, T., Ogino, S., Kirchner, T., and Shivdasani, R. a. (2012). Differential WNT activity in colorectal cancer confers limited tumorigenic potential and is regulated by MAPK signaling. Cancer Res. *72*, 1547–1556.

Huang, J., Nguyen-McCarty, M., Hexner, E.O., Danet-Desnoyers, G., and Klein, P.S. (2012a). Maintenance of hematopoietic stem cells through regulation of Wnt and mTOR pathways. Nat. Med. *18*, 1778–1785.

Huang, R.Y., Guilford, P., Thiery, J.P., and Shroom, R. (2012b). Early Events in Cell Adhesion and polarity during Epithelial-Mesenchymal Transition. J. Cell Sci. *125*, 4417-22.

Humphries, A., and Wright, N. a (2008). Colonic crypt organization and tumorigenesis. Nat. Rev. Cancer *8*, 415–424.

Humphries, A., Cereser, B., Gay, L.J., Miller, D.S.J., Das, B., Gutteridge, A., Elia, G., Nye, E., Jeffery, R., Poulsom, R., et al. (2013). Lineage tracing reveals multipotent stem cells maintain human adenomas and the pattern of clonal expansion in tumor evolution. Proc. Natl. Acad. Sci. *110*, E2490–E2499.

Huntly, B.J.P., and Gilliland, D.G. (2005). Leukaemia stem cells and the evolution of cancer-stem-cell research. Nat. Rev. Cancer *5*, 311–321.

Itzkovitz, S., Lyubimova, A., Blat, I., Maynard, M., Es, J. Van, Lees, J., Jacks, T., Clevers, H., and Oudenaarden, A. Van (2012). Single molecule transcript counting of stem cell markers in the mouse intestine. Nat Cell Biol. *14*, 106–114.

Iwano, M., Plieth, D., Danoff, T.M., Xue, C., Okada, H., and Neilson, E.G. (2002). Evidence that fibroblasts derive from epithelium during tissue fibrosis. J. Clin. Invest. *110*, 341–350.

Jackstadt, R., Röh, S., Neumann, J., Jung, P., Hoffmann, R., Horst, D., Berens, C., Bornkamm, G.W., Kirchner, T., Menssen, A., et al. (2013). AP4 is a mediator of epithelial-mesenchymal transition and metastasis in colorectal cancer. J. Exp. Med. *210*, 1331–1350.

Jackstadt, R., Hermeking, H., Jackstadt, R., and Hermeking, H. (2014). AP4 is required for mitogen- and c-MYC-induced cell cycle progression. Oncotarget *5*, 7316–7327.

Jass, J.R., Young, J., and Leggett, B.A. (2002). Evolution of colorectal cancer: Change of pace and change of direction. J. Gastroenterol. Hepatol. *17*, 17–26.

Jemal, A., Siegel, R., Ward, E., Hao, Y., Xu, J., and Thun, M.J. (2009). Cancer statistics, 2009. CA. Cancer J. Clin. *59*, 225–249.

Kahn, M. (2014). Kahn, M. (2014). Can we safely target the WNT pathway? Nature Reviews. Drug Discovery, 13(7), 513–32. doi:10.1038/nrd4233Can we safely target the WNT pathway? Nat. Rev. Drug Discov. *13*, 513–532.

Kaller, M., Liffers, S.-T., Oeljeklaus, S., Kuhlmann, K., Roh, S., Hoffmann, R., Warscheid, B., and Hermeking, H. (2011). Genome-wide Characterization of miR-34a Induced Changes in Protein and mRNA Expression by a Combined Pulsed SILAC and Microarray Analysis. Mol. Cell. Proteomics *10*, M111.010462-M111.010462.

Karimi, P., Islami, F., Anandasabapathy, S., Freedman, N.D., and Kamangar, F. (2014). Gastric Cancer: Descriptive Epidemiology, Risk Factors, Screening, and Prevention. Cancer Epidemiol. Biomarkers Prev. *23*, 700–713.

Kikugawa, T., Kinugasa, Y., Shiraishi, K., Nanba, D., Nakashiro, K.I., Tanji, N., Yokoyama, M., and Higashiyama, S. (2006). PLZF regulates Pbx1 transcription and Pbx1-HoxC8 complex leads to androgen-independent prostate cancer proliferation. Prostate *66*, 1092–1099.

Kirchner, T., and Brabletz, T. (2000). Patterning and nuclear beta-catenin expression in the colonic adenoma-carcinoma sequence. Analogies with embryonic gastrulation. Am. J. Pathol. *157*, 1113–1121.

Korinek, V., Barker, N., Moerer, P., van Donselaar, E., Huls, G., Peters, P.J., and Clevers, H. (1998). Depletion of epithelial stem-cell compartments in the small intestine of mice lacking Tcf-4. Nat. Genet. *19*, 379–383.

Korolev, K.S., Avlund, M., Hallatschek, O., and Nelson, D.R. (2010). Genetic demixing and evolution in linear stepping stone models. Rev. Mod. Phys. *82*, 1691–1718.

Kreso, A., and Dick, J.E. (2014). Evolution of the cancer stem cell model. Cell Stem Cell *14*, 275–291.

Lao, V.V., and Grady, W.M. (2011). Epigenetics and colorectal cancer. Nat. Rev. Gastroenterol. Hepatol. *8*, 686–700.

Lathia, J.D., Hitomi, M., Gallagher, J., Gadani, S.P., Adkins, J., Vasanji, A., Liu, L., Eyler, C.E., Heddleston, J.M., Wu, Q., et al. (2011). Distribution of CD133 reveals glioma stem cells self-renew through symmetric and asymmetric cell divisions. Cell Death Dis *2*, e200.

Lewis, B.P., Burge, C.B., and Bartel, D.P. (2005). Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. Cell *120*, 15–20.

Li, B., Zhang, S., Shen, H., and Li, C. (2017). MicroRNA-144-3p suppresses gastric cancer progression by inhibiting epithelial-to-mesenchymal transition through targeting PBX3. Biochem. Biophys. Res. Commun. *484*, 241–247.

Li, V.S.W., Ng, S.S., Boersema, P.J., Low, T.Y., Karthaus, W.R., Gerlach, J.P., Mohammed, S., Heck, A.J.R., Maurice, M.M., Mahmoudi, T., et al. (2012). Wnt Signaling through Inhibition of β-Catenin Degradation in an Intact Axin1 Complex. Cell *149*, 1245–1256.

Li, Z., Zhang, Z., Li, Y., Arnovitz, S., Chen, P., Huang, H., Jiang, X., Hong, G.M., Kunjamma, R.B., Ren, H., et al. (2013). PBX3 is an important cofactor of HOXA9 in leukemogenesis. Blood *121*, 1422–1431.

Liu, C., Li, Y., Semenov, M., Han, C., Baeg, G.H., Tan, Y., Zhang, Z., Lin, X., and He, X. (2002). Control of β-catenin phosphorylation/degradation by a dual-kinase mechanism. Cell *108*, 837–847.

Livet, J., Weissman, T. a, Kang, H., Draft, R.W., Lu, J., Bennis, R. a, Sanes, J.R., and Lichtman, J.W. (2007). Transgenic strategies for combinatorial expression of fluorescent proteins in the nervous system. Nature *450*, 56–62.

Logan, C.Y., and Nusse, R. (2004). THE WNT SIGNALING PATHWAY IN DEVELOPMENT AND DISEASE. Annu. Rev. Cell Dev. Biol. *20*, 781–810.

Lopez-Garcia, C., Klein, A.M., Simons, B.D., and Winton, D.J. (2010). Intestinal stem cell replacement follows a pattern of neutral drift. Science (80). *330*, 822–825.

Macías-Silva, M., Abdollah, S., Hoodless, P. a, Pirone, R., Attisano, L., and Wrana, J.L. (1996). MADR2 is a substrate of the TGFbeta receptor and its phosphorylation is required for nuclear accumulation and signaling. Cell *87*, 1215–1224.

Mani, S.A., Yang, J., Brooks, M., Schwaninger, G., Zhou, A., Miura, N., Kutok, J.L., Hartwell, K., Richardson, A.L., and Weinberg, R.A. (2007). Mesenchyme Forkhead 1 (FOXC2) plays a key role in metastasis and is associated with aggressive basal-like breast cancers. Proc. Natl. Acad. Sci. U. S. A. *104*, 10069–10074.

Matsuda, T., and Cepko, C.L. (2007). Controlled expression of transgenes introduced by in vivo electroporation. Proc. Natl. Acad. Sci. (U.S.A.) *104*, 1027–1032.

McFarland, C.D., Korolev, K.S., Kryukov, G. V, Sunyaev, S.R., and Mirny, L.A. (2013). Impact of deleterious passenger mutations on cancer progression. Proc. Natl. Acad. Sci. U. S. A. *110*, 2910–2915.

Meacham, C.E., and Morrison, S.J. (2013). Tumour heterogeneity and cancer cell plasticity. Nature *501*, 328–337.

Michor, F., and Polyak, K. (2010). The origins and implications of intratumor heterogeneity. Cancer Prev. Res. *3*, 1361–1364.

Miller, J.R. (2002). The Wnts. Genome Biol. *3*, REVIEWS3001.

Monica, K., Galili, N., Nourse, J., Saltman, D., and Cleary, M.L. (1991). PBX2 and PBX3, new homeobox genes with extensive homology to the human proto-oncogene PBX1. Mol. Cell. Biol. *11*, 6149–6157.

Montgomery, R.K., Carlone, D.L., Richmond, C. a, Farilla, L., Kranendonk, M.E.G., Henderson, D.E., Baffour-Awuah, N.Y., Ambruzs, D.M., Fogli, L.K., Algra, S., et al. (2011). Mouse telomerase reverse transcriptase (mTert) expression marks slowly cycling intestinal stem cells. Proc. Natl. Acad. Sci. U. S. A. *108*, 179–184.

Moreno-Bueno, G., Peinado, H., Molina, P., Olmeda, D., Cubillo, E., Santos, V., Palacios, J., Portillo, F., and Cano, A. (2009). The morphological and molecular features of the epithelial-to-mesenchymal transition. Nat Protoc *4*, 1591–1613.

Muñoz, J., Stange, D.E., Schepers, A.G., van de Wetering, M., Koo, B.-K., Itzkovitz, S., Volckmann, R., Kung, K.S., Koster, J., Radulescu, S., et al. (2012). The Lgr5 intestinal stem cell signature: robust expression of proposed quiescent “+4” cell markers. EMBO J. *31*, 3079–3091.

Nieto, M.A. (2013). Epithelial plasticity: a common theme in embryonic and cancer cells. Science (80). *342*, 1234850.

Nieto, M. Angela, Huang, R.Y.J., Jackson, R.A., and Thiery, J.P. (2016). EMT: 2016. Cell *166*, 21–45.

Nosho, K., Irahara, N., Shima, K., Kure, S., Kirkner, G.J., Schernhammer, E.S., Hazra, A., Hunter, D.J., Quackenbush, J., Spiegelman, D., et al. (2008). Comprehensive biostatistical analysis of CpG island methylator phenotype in colorectal cancer using a large population-based sample. PLoS One *3*, e3698.

Oliver E. Owen, S.C.K. and R.W.H. (2002). J. Biol. Chem.-2010-Anthonisen-1607-15.pdf. J. Biol. Chem. *277*, 30409–30412.

Orth, M.F., Cazes, A., Butt, E., and Grunewald, T.G.P. (2015). An update on the LIM and SH3 domain protein 1 (LASP1): a versatile structural, signaling, and biomarker protein. Oncotarget *6*, 26–42.

Peinado, H., Olmeda, D., and Cano, A. (2007). Snail, Zeb and bHLH factors in tumour progression: an alliance against the epithelial phenotype? Nat. Rev. Cancer *7*, 415–428.

Pérez-Moreno, M.A., Locascio, A., Rodrigo, I., Dhondt, G., Portillo, F., Nieto, M.A., and Cano, A. (2001). A New Role for E12/E47 in the Repression of E-cadherin Expression and Epithelial-Mesenchymal Transitions. J. Biol. Chem. *276*, 27424–27431.

Pillai, R.S. (2005). Inhibition of Translational Initiation by Let-7 MicroRNA in Human Cells. Science (80). *309*, 1573–1576.

Plaks, V., Kong, N., and Werb, Z. (2015). The cancer stem cell niche: How essential is the niche in regulating stemness of tumor cells? Cell Stem Cell *16*, 225–238.

Ponder, B.A., Schmidt, G.H., Wilkinson, M.M., Wood, M.J., Monk, M., and Reid, A. (1985). Derivation of mouse intestinal crypts from single progenitor cells. Nature *313*, 689–691.

Potten, C.S., Wilson, J.W., and Booth, C. (1997). Regulation and significance of apoptosis in the stem cells of the gastrointestinal epithelium. Stem Cells *15*, 82–93.

Ramberg, H., Alshbib, A., Berge, V., Svindland, A., and Taskén, K.A. (2011). Regulation of PBX3 expression by androgen and Let-7d in prostate cancer. Mol. Cancer *10*, 50.

Reya, T., Morrison, S.J., Clarke, M.F., and Weissman, I.L. (2001). Stem cells, cancer, and cancer stem cells. Nature *414*, 105.

Sahay, D., Leblanc, R., Grunewald, T.G.P., Ambatipudi, S., Ribeiro, J., Clézardin, P., and Peyruchaud, O. (2015). The LPA1/ZEB1/miR-21-activation pathway regulates metastasis in basal breast cancer. Oncotarget *6*, 20604–20620.

Samowitz, W.S., Albertsen, H., Herrick, J., Levin, T.R., Sweeney, C., Murtaugh, M.A., Wolff, R.K., and Slattery, M.L. (2005). Evaluation of a large, population-based sample supports a CpG island methylator phenotype in colon cancer. Gastroenterology *129*, 837–845.

Sanchez-Tillo, E., de Barrios, O., Siles, L., Cuatrecasas, M., Castells, A., and Postigo, A. (2011). β-catenin/TCF4 complex induces the epithelial-to-mesenchymal transition (EMT)-activator ZEB1 to regulate tumor invasiveness. Proc. Natl. Acad. Sci. *108*, 19204–19209.

Sánchez-Tilló, E., Liu, Y., De Barrios, O., Siles, L., Fanlo, L., Cuatrecasas, M., Darling, D.S., Dean, D.C., Castells, A., and Postigo, A. (2012). EMT-activating transcription factors in cancer: Beyond EMT and tumor invasiveness. Cell. Mol. Life Sci. *69*, 3429–3456.

Schepers, A.G., Snippert, H.J., Stange, D.E., Born, M. van den, Es, J.H. van, Wetering, M. van de, and Clevers, H. (2012). Lineage Tracing Reveals Lgr5+ Stem Cell Activity in Mouse Intestinal Adenomas. Science (80). *337*, 730–735.

Shah, S.P., Morin, R.D., Khattra, J., Prentice, L., Pugh, T., Burleigh, A., Delaney, A., Gelmon, K., Guliany, R., Senz, J., et al. (2009). Mutational evolution in a lobular breast tumour profiled at single nucleotide resolution. Nature *461*, 809–813.

Shimokawa, M., Ohta, Y., Nishikori, S., Matano, M., Takano, A., Fujii, M., Date, S., Sugimoto, S., Kanai, T., and Sato, T. (2017). Visualization and targeting of LGR5+ human colon cancer stem cells. Nature *545*, 187-192.

Shiraishi, K., Yamasaki, K., Nanba, D., Inoue, H., Hanakawa, Y., Shirakata, Y., Hashimoto, K., and Higashiyama, S. (2007). Pre-B-cell leukemia transcription factor 1 is a major target of promyelocytic leukemia zinc-finger-mediated melanoma cell growth suppression. Oncogene *26*, 339–348.

Siegmund, K.D., Marjoram, P., Woo, Y.-J., Tavaré, S., and Shibata, D. (2009). Inferring clonal expansion and cancer stem cell dynamics from DNA methylation patterns in colorectal cancers. Proc. Natl. Acad. Sci. U. S. A. *106*, 4828–4833.

Siemens, H., Jackstadt, R., Hünten, S., Kaller, M., Menssen, A., Götz, U., and Hermeking, H. (2011). miR-34 and SNAIL form a double-negative feedback loop to regulate epithelial-mesenchymal transitions. Cell Cycle *10*, 4256–4271.

Sjöblom, T., Jones, S., Wood, L.D., Parsons, D.W., Lin, J., Barber, T.D., Mandelker, D., Leary, R.J., Ptak, J., Silliman, N., et al. (2006). The consensus coding sequences of human breast and colorectal cancers. Science *314*, 268–274.

Snippert, H.J., van der Flier, L.G., Sato, T., van Es, J.H., van den Born, M., Kroon-Veenboer, C., Barker, N., Klein, A.M., van Rheenen, J., Simons, B.D., et al. (2010). Intestinal crypt homeostasis results from neutral competition between symmetrically dividing Lgr5 stem cells. Cell *143*, 134–144.

Sottoriva, A., Kang, H., Ma, Z., Graham, T.A., Salomon, M.P., Zhao, J., Marjoram, P., Siegmund, K., Press, M.F., Shibata, D., et al. (2015). A Big Bang model of human colorectal tumor growth. Nat. Genet. *47*, 209–216.

Spaderna, S., Schmalhofer, O., Hlubek, F., Berx, G., Eger, A., Merkel, S., Jung, A., Kirchner, T., and Brabletz, T. (2006). A Transient, EMT-Linked Loss of Basement Membranes Indicates Metastasis and Poor Survival in Colorectal Cancer. Gastroenterology *131*, 830–840.

Staal, F.J., and Clevers, H.C. (2005). WNT signalling and haematopoiesis: a WNT-WNT situation. Nat Rev Immunol *5*, 21–30.

Takaku, K., Oshima, M., Miyoshi, H., Matsui, M., Seldin, M.F., and Taketo, M.M. (1998). Intestinal Tumorigenesis in Compound Mutant Mice of both Dpc4(Smad4) and Apc Genes. Cell *92*, 645–656.

Takeda, N., Jain, R., LeBoeuf, M.R., Wang, Q., Lu, M.M., and Epstein, J.A. (2011). Interconversion between intestinal stem cell populations in distinct niches. Science *334*, 1420–1424.

Tam, W.L., and Weinberg, R.A. (2013). The epigenetics of epithelial-mesenchymal plasticity in cancer. Nat. Med. *19*, 1438–1449.

Taube, J.H., Herschkowitz, J.I., Komurov, K., Zhou, A.Y., Gupta, S., Yang, J., Hartwell, K., Onder, T.T., Gupta, P.B., Evans, K.W., et al. (2010). Core epithelial-to-mesenchymal transition interactome gene-expression signature is associated with claudin-low and metaplastic breast cancer subtypes. Proc. Natl. Acad. Sci. *107*, 15449–15454.

Tetsu, O., and McCormick, F. (1999). Beta-catenin regulates expression of cyclin D1 in colon carcinoma cells. Nature *398*, 422–426.

The Cancer Genome Network Atlas (2012). Comprehensive molecular characterization of human colon and rectal cancer. Nature *487*, 330–337.

Thiery, J.P., and Sleeman, J.P. (2006). Complex networks orchestrate epithelial-mesenchymal transitions. Nat. Rev. Mol. Cell Biol. *7*, 131–142.

Thiery, J.P., Acloque, H., Huang, R.Y.J., and Nieto, M.A. (2009). Epithelial-Mesenchymal Transitions in Development and Disease. Cell *139*, 871–890.

Toiyama, Y., Yasuda, H., Saigusa, S., Tanaka, K., Inoue, Y., Goel, A., and Kusunoki, M. (2013). Increased expression of slug and vimentin as novel predictive biomarkers for lymph node metastasis and poor prognosis in colorectal cancer. Carcinogenesis *34*, 2548–2557.

Torre, L.A., Bray, F., Siegel, R.L., Ferlay, J., Lortet-Tieulent, J., and Jemal, A. (2015). Global cancer statistics, 2012. CA. Cancer J. Clin. *65*, 87–108.

Vazquez, A., Bond, E.E., Levine, A.J., and Bond, G.L. (2008). The genetics of the p53 pathway, apoptosis and cancer therapy. Nat. Rev. Drug Discov. *7*, 979–987.

Vermeulen, L., and Snippert, H.J. (2014). Stem cell dynamics in homeostasis and cancer of the intestine. Nat. Rev. Cancer *14*, 468–480.

Vermeulen, L., De Sousa E Melo, F., van der Heijden, M., Cameron, K., de Jong, J.H., Borovski, T., Tuynman, J.B., Todaro, M., Merz, C., Rodermond, H., et al. (2010). Wnt activity defines colon cancer stem cells and is regulated by the microenvironment. Nat. Cell Biol. *12*, 468–476.

Vial, E., Sahai, E., and Marshall, C.J. (2003). ERK-MAPK signaling coordinately regulates activity of Rac1 and RhoA for tumor cell motility. Cancer Cell *4*, 67–79.

Vogelstein, B., Fearon, E.R., Hamilton, S.R., Kern, S.E., Preisinger, A.C., Leppert, M., Nakamura, Y., White, R., Smits, A.M.M., and Bos, J.L. (1988). Genetic Alterations during Colorectal-Tumor Development. N. Engl. J. Med. *319*, 525–532.

Vogelstein, B., Papadopoulos, N., Velculescu, V.E., Zhou, S., Diaz Jr., L.A., and Kinzler, K.W. (2013). Cancer Genome Landscapes. Science (80). *339*, 1546–1558.

Wang, X., Zheng, M., Liu, G., Xia, W., McKeown-Longo, P.J., Hung, M.C., and Zhao, J. (2007). Kruppel-like factor 8 induces epithelial to mesenchymal transition and epithelial cell invasion. Cancer Res *67*, 7184–7193.

Wang, Y.K., Zhu, Y.L., Qiu, F.M., Zhang, T., Chen, Z.G., Zheng, S., and Huang, J. (2010). Activation of Akt and MAPK pathways enhances the tumorigenicity of CD133+ primary colon cancer cells. Carcinogenesis *31*, 1376–1380.

Weber, K., Thomaschewski, M., Warlich, M., Volz, T., Cornils, K., Niebuhr, B., Täger, M., Lütgehetmann, M., Pollok, J.-M.M., Stocking, C., et al. (2011). RGB marking facilitates multicolor clonal cell tracking. Nat. Med. *17*, 504–509.

Welch, C., Chen, Y., and Stallings, R.L. (2007). MicroRNA-34a functions as a potential tumor suppressor by inducing apoptosis in neuroblastoma cells. Oncogene *26*, 5017–5022.

Van de Wetering, M., Sancho, E., Verweij, C., De Lau, W., Oving, I., Hurlstone, A., Van der Horn, K., Batlle, E., Coudreuse, D., Haramis, A.P., et al. (2002). The β-catenin/TCF-4 complex imposes a crypt progenitor phenotype on colorectal cancer cells. Cell *111*, 241–250.

Wheelock, M.J., Shintani, Y., Maeda, M., Fukumoto, Y., and Johnson, K.R. (2008). Cadherin switching. J. Cell Sci. *121*, 727–735.

Wodarz, a, and Nusse, R. (1998). Mechanisms of Wnt signaling in development. Annu. Rev. Cell Dev. Biol. *14*, 59–88.

Wood, L.D., Parsons, D.W., Jones, S., Lin, J., Sjöblom, T., Leary, R.J., Shen, D., Boca, S.M., Barber, T., Ptak, J., et al. (2007). The genomic landscapes of human breast and colorectal cancers. Science *318*, 1108–1113.

World Cancer Research Fund, A.I. for C.R. (2007). Food, Nutrition, Physical Activity, and the Prevention of Cancer: a Global Perspective. Washingt. DC AICR.

Yan, D., Wiesmann, M., Rohan, M., Chan, V., Jefferson, A.B., Guo, L., Sakamoto, D., Caothien, R.H., Fuller, J.H., Reinhard, C., et al. (2001). Elevated expression of axin2 and hnkd mRNA provides evidence that Wnt/β-catenin signaling is activated in human colon tumors. Proc. Natl. Acad. Sci. U.

S. A. *98*, 14973–14978.

Yanagawa, S.I., Matsuda, Y., Lee, J.S., Matsubayashi, H., Sese, S., Kadowaki, T., and Ishimoto, A. (2002). Casein kinase I phosphorylates the Armadillo protein and induces its degradation in Drosophila. EMBO J. *21*, 1733–1742.

Yang, J., Mani, S.A., Donaher, J.L., Ramaswamy, S., Itzykson, R.A., Come, C., Savagner, P., Gitelman, I., Richardson, A., and Weinberg, R.A. (2004). Twist, a master regulator of morphogenesis, plays an essential role in tumor metastasis. Cell *117*, 927–939.

Yook, J.I., Li, X.-Y., Ota, I., Hu, C., Kim, H.S., Kim, N.H., Cha, S.Y., Ryu, J.K., Choi, Y.J., Kim, J., et al. (2006). A Wnt-Axin2-GSK3beta cascade regulates Snail1 activity in breast cancer cells. Nat. Cell Biol. *8*, 1398–1406.

Yost, C., Torres, M., Miller, J.R., Huang, E., Kimelman, D., and Moon, R.T. (1996). The axis-inducing activity, stability, and subcellular distribution of beta-catenin is regulated in Xenopus embryos by glycogen synthase kinase 3. Genes Dev. *10*, 1443–1454.

Zhou, B.P., Deng, J., Xia, W., Xu, J., Li, Y.M., Gunduz, M., and Hung, M.-C. (2004). Dual regulation of Snail by GSK-3beta-mediated phosphorylation in control of epithelial-mesenchymal transition. Nat. Cell Biol. *6*, 931–940.

Zomer, A., Ellenbroek, S.I.J., Ritsma, L., Beerling, E., Vrisekoop, N., and Van Rheenen, J. (2013). Brief report: Intravital imaging of cancer stem cell plasticity in mammary tumors. Stem Cells *31*, 602–606.

Zufferey, R., Nagy, D., Mandel, R.J., Naldini, L., and Trono, D. (1997). Multiply attenuated lentiviral vector achieves efficient gene delivery in vivo. Nat. Biotechnol. *15*, 871–875.